

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM Protein - protein search, using sw model

Run on: August 10, 2004, 15:18:30 ; Search time 84.6574 Seconds

(without alignments)
1632.058 Million cell updates/secTitle: US-09-455-978B-2
Perfect score: 23.94
Sequence: 1 MSNDNTLTVADVRNGIDGH.....ATDQQVRTVEEVRETVGKLS 489Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:
1586107Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 10%
Listing First 45 summaries

Database :

- A_Geneseq_29Jan04:*
- 1: Geneseq_1905:*
- 2: Geneseq_1990s:*
- 3: Geneseq_2000s:*
- 4: Geneseq_2001s:*
- 5: Geneseq_2002s:*
- 6: Geneseq_2003as:*
- 7: Geneseq_2003bs:*
- 8: Geneseq_2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	23.94	100.0	489	4	AAE04641	Aae0641 Halobacte
2	41.2.5	17.2	432	4	AAE04642	Aae0642 Bacillus
3	399.5	16.7	535	6	ABU38469	Protein e
4	387.5	16.2	739	4	ABAB96493	Putative
5	386.5	16.1	682	6	Abu38230	Protein e
6	380	15.9	435	4	ABAB96483	Putative
7	375	15.7	501	4	ABAB95709	Putative
8	374.5	15.6	680	6	ABU41746	Protein e
9	373.5	15.6	709	6	ADA33885	Acinetoba
10	371.5	15.5	686	6	ABU40238	Protein e
11	366	15.1	644	6	ABU49667	Protein e
12	359.5	15.0	891	3	ABAB26596	Synechocystis
13	356.5	14.9	374	4	ABAB95382	Putative
14	356.5	14.9	539	6	ABU23101	Protein e
15	348	14.5	626	6	ABU49142	Protein e
16	346.5	14.5	565	6	ABM68096	Photorhab
17	346.5	14.2	601	5	ABBA4909	Listeria
18	338	14.1	686	6	ABU24381	Protein e
19	326	13.6	545	6	ABM68095	Photorhab
20	326	13.6	641	6	ABU49354	Protein e
21	315	13.2	643	6	ABU492678	Protein e
22	314.5	13.1	501	6	ABU41640	Protein e
23	313.5	13.1	541	6	ABU49123	Protein e
24	308	12.9	845	6	ABU48533	Protein e
25	298	12.4	511	3	ABAB14148	Bordetell

ALIGNMENTS

RESULT 1	AAE04641	ID AAE04641 standard; protein; 489 AA.
XX	XX	XX
AC	AC	AC
XX	XX	XX
DT	11-SEP-2003 (revised)	DT 04-SEP-2001 (first entry)
XX	XX	XX
DE	Halobacterium salinarum HemAT-Hs protein.	KW
XX	XX	KW Haem binding protein; HemAT-Hs; HemAT-Bs; gaseous ligand sensor; oxygen storage; artificial photosynthesis; signalling function; alpha-haemoglobin; myoglobin; therapy.
XX	OS	KW alpha-haemoglobin; myoglobin; therapy.
XX	OS	KW Halobacterium salinarum.
XX	OS	XX
PN	WO200140475-A2.	PN WO200140475-A2.
XX	XX	XX
PD	07-JUN-2001.	PD 07-JUN-2001.
XX	XX	XX
PF	2000WO-US033048.	PF 05-DEC-2000; 2000WO-US033048.
XX	XX	PR 06-DEC-1999;
XX	XX	PA 99US-00455978.
XX	XX	PA (UYHA-) UNIV HAWAII.
XX	XX	PI Alam M, Larsen R;
XX	XX	XX DR WPI; 2001-374832/39.
XX	XX	DR N-PSDB; AAD08991.
XX	XX	PS Claim 6; Page 10; 94pp; English.
XX	XX	XX The present invention relates to isolated archael and bacterial haem binding protein, HemAT-Hs or HemAT-Bs which reversibly binds oxygen with low affinity. Haem binding protein is useful for controlled storage of oxygen by allowing haem binding protein to bind and store oxygen, and triggering the release of oxygen from haem binding protein by activating the signalling domain. Haem binding protein is useful for sensing gaseous ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem binding protein is useful for treating a patient suffering from low blood levels by administering and regulating the oxygen binding of the haem binding protein by modifying the signalling domain. Haem binding protein

CC is useful for haem-based catalysis, for artificial photosynthesis and for identifying potential signalling functions of mutated alpha haemoglobin and myoglobin causing several diseases. The present sequence is CC Halobacterium salinarium HemAT-Hs protein which is salt tolerant. CC (Updated on 11-SEP-2003 to standardise OS field)

SQ Sequence 489 AA;

Query Match 100.0%; Score 2394; DB 4; Length 489;
Best Local Similarity 100.0%; Pred. No. 6.3e-165; Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 489; Sequence 432 AA;

PA (UHFA-) UNIV HAWAII.
XX Alam M, Larsen R;
PI XX
WPI: 2001-374832/39.
DR N-PSDB; AAD08932.
XX

PT Novel isolated bacterial heme binding protein, HemAT-Hs or HemAT-Bs which reversibly binds oxygen with low affinity, useful for controlled storage of oxygen and for sensing gaseous ligands such as oxygen.

Claim 8; Page 11; 94pp; English.

The present invention relates to isolated archael and bacterial haem binding protein, HemAT-Hs or HemAT-Bs which reversibly binds oxygen with low affinity. Haem binding protein is useful for controlled storage of oxygen by allowing haem binding protein to bind and store oxygen, and triggering the release of oxygen from haem binding protein by activating the signalling domain. Haem binding protein is useful for sensing gaseous ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem binding protein is useful for treating a patient suffering from low blood levels by administering and regulating the oxygen binding of the haem binding protein by modifying the signalling domain. Haem binding protein is useful for haem-based catalysis, for artificial photosynthesis and for identifying potential signalling functions of mutated alpha haemoglobin and myoglobin causing several diseases. The present sequence is Bacillus subtilis HemAT-Bs protein.

Sequence 432 AA;

Query Match 17.2%; Score 412.5; DB 4; Length 432;
Best Local Similarity 26.2%; Pred. No. 1.4e-21;
Matches 114; Conservative 85; Mismatches 187; Indels 49; Gaps 8;

Qy 31 AEIAWILSFTGIDDTMAALAAEOPLFET 60
Db 32 ADVKQLKMYVLGDAALYVLEQLOPLIQLQENIVNIDAFYKNLDHESLMDII-NDHSVVD 90

Qy 91 QLKETQAEYLGLGRGEYDTEYAAQARIGKTHDVILGLPDPVLYGATRYTGTGLDALAD 150
Db 91 RLKQTLLKRHQEMFAQVDPDEIEKENRIASIHRLIGLPPWYGAQELLSMIDY-- 148

Qy 151 DVVADRGEEAAAVIDELVARELPMLKLTFDQIQAMTDYDSYAQLRHEDDSRQEELANA 210
Db 149 -----EASTINQOELLKAIAKATVILNLBQOLVLEAQSEYNO-TRDQEKEKNLH- 199

Qy 211 VATHVEAPLSSLEATSDQDVARTDMRARTDDQVDRMADSREISSVSVASBEVASTADD 270
Db 200 -----QKIQETSGSIA---NLFSEIERSVQELVDSEGISQSASKAGTVTSVSEE 246

Qy 271 VRRTSEDAEAQAOQGEAAADDALATMTDIDEATGVTAGYVQGLGERAADYBSVTGVYDDI 330
Db 247 -KSIGGKKELEVQQKQ-----MNKIDTSLVQIEKEMVKLDDEAQQIEKIFGIVGTI 296

Qy 331 AEQTNNMIALNASTEARAGEAGEGFAVVADEVKALAAESREQSTREVELYEQMQAETET 390
Db 297 AEQTNNLISLNASTEARAGEGKFAVVADEVKALAAESREQSTREVELYEQMQAETET 356

Qy 391 VDQLEVNQRIGEGVERVEEAMETLQBIDAVEDA-----ASGMQEVSTAT 436
Db 357 SKHIDVNLVSEKSMKTOINRFDEIVSMSMKISKEQSCKIDVLDLQAFLGGQEVYRSRAV 416

Qy 437 DEQAVSTEVAEMVD 451
Db 417 SHVAASVDSLVILTE 431

RESULT 3
ABU38469

PD 07-JUN-2001.
XX 05-DEC-2000; 2000WO-US033048.
XX 06-DEC-1999; 99US-00455978.
XX AC ABU38469;
XX

AC AAE04642;
XX DT 04-SEP-2001 (first entry)
XX DE Bacillus subtilis HemAT-Bs protein.
XX KW Haem binding protein; HemAT-Hs; HemAT-Bs; gaseous ligand sensor;
KW oxygen storage; artificial photosynthesis; signalling function;
KW alpha haemoglobin; myoglobin; therapy.
XX OS Bacillus subtilis.
PN WO200140475-A2.
XX PD 07-JUN-2001.
XX KW Haem binding protein; HemAT-Hs; HemAT-Bs; gaseous ligand sensor;
KW oxygen storage; artificial photosynthesis; signalling function;
KW alpha haemoglobin; myoglobin; therapy.
XX OS Bacillus subtilis.
PN WO200140475-A2.
XX AC ABU38469;
XX

CC a hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal vents. The present sequence is one such *P. abyssi* protein. CC The proteins of the present invention have various potential industrial uses, since the proteins are stable at very high temperatures, some up to CC 110 degrees centigrade. Note: This patent is in the same patent family as CC WO20060206, which contains additional sequences as shown in AAB99132- CC AAB99143, AAH75903-AAH75920 and Aad66436.

XX Sequence 739 AA;
SQ Query Match 16.2%; Score 387.5; DB 4; Length 739;
Best Local Similarity 28.6%; Pred. No. 1.8e-19;
Matches 126; Conservative 83; Mismatches 150; Indels 81; Gaps 15;

Db 83 ANSTKTEQKETQAEWILIGLQRGEYDTETIAQ---RARIGKI----HDVFLGLG 129
Db 331 SNTLAPLEKLK--YAAQALAEGRLLKQSYEYLKQIRYLERDIGALIQAFEAWSKOLVGT- 387
Qy 130 PDVYLGAYTRYTYGLDALLAD-----VVADRGEAAAAYDELVARFLPMLKLUFTDQQ 183
Db 388 ---INRISK---KLERLAEEDLSNLTVTVERGE-----LRDIED---421
Db 422 ----IRSVTETFREISGLSLEYMAN---DIEKRANALAQSKDVTEAINQVN-PAIQQ 470
Qy 184 IAMDTYTSYAAQLRHEIDSROELANDAVATHVEAPLSLEATSODVAERTDMARTDDQ 243
Qy 471 VSIAQRQRQETINEITDGMRLLVATQCTSEESVRAAMEEFSGAVTEVVSIANGSQKDEALKR 530
Db 296 MTDIDEATGVTAGVTDQGEVQGLERAADVESVTGVDDIAREQTMNLNASIEARAGEAGBGF 355
Db 531 FEDICHMMMSRLFETYSKVAAENSRNIEEITINVITSIEEQTNLMLNARATEBARAGRGF 590
Qy 356 AVVADEVKALAESREQSTREVLELVEOMQAETEVTVDQDLEVNRQIGVERVEEAMETL 415
Db 591 AVVAQEIRKLAKESQKAQNPKSIIKDITDEIKEAVATK-----EGTGSVIGESETL 643
Qy 416 QEIIDAVEADAASGMQEVS---TATDQAVQSTREVAE---MDVGDVDRAGEIAAALDDIA 468
Db 644 RDTGIGLNIATLQLQESTSERMTVKECIOVTRTOBEVDKALRALENLAASAEETTASAEEVS 703
Qy 469 DATDQQVRTVEEVRETGYKL 488
Db 704 SAIEQTATAEFLRRAAQEL 723

XX Sequence 682 AA;

Query Match 16.1%; Score 386.5; DB 6; Length 682;
Best Local Similarity 27.5%; Pred. No. 1.9e-19;
Matches 95; Conservative 77; Mismatches 140; Indels 33; Gaps 4;
Qy 144 LLDAADDWVADRGEEAAAVDELVARFLPMKLTFDQQIAMDDYDSYAQRLHEDIDS 203
Db 356 LLDEIAD--LADGDTVAAATYE-----DFTGTAADSINYSDQ 392
Qy 204 ROELA---NAVATHVEAPLSSLEATSQDVAAERTDTMARTDDQVDRMADSREISSVSAS 260
Db 393 LRELVETINQTAQVQVAAAQETQSFMHLAE-----ASEHQAEQIAQGSAAINEMAVS 445
Qy 261 VEEVASTADDVVRTSSEDAEAIAQGCEAAADDALATMIDIDEATDGVTTAGVEQLGERAADV 320
Db 446 IDQVANASESSAVALRSVALANKNEVVNTITGNDNIREQIQTSKRKLGRSSQE 505

XX Sequence 682 AA;

Query Match 16.1%; Score 386.5; DB 6; Length 682;
Best Local Similarity 27.5%; Pred. No. 1.9e-19;
Matches 95; Conservative 77; Mismatches 140; Indels 33; Gaps 4;
Qy 321 ESVTGYYDDIAEQTQNMALANASIEAARAGEAGEFIAVVADEVKALAAEESREQSTIVEELY 380
Db 506 GDIVSLINDLADQTMNLNALNAAIQSMAGDAGRGGFIAVVADEVQRLAERSAAATKQIEALV 565
Qy 381 EQMQAEETEYTVDQDDEVNQRIGEYGRVEEAMETLQBTDAVEDAASGMQEVSTADEQA 440
Db 566 KTIQDTNEAVISMEQTTSEVVRGARLAQDGALEBIEKVKSTLAQIIONSNAAARQQA 625
Qy 441 VSTEVEVAEMVGDVDDRDGETAAALDDIADATDQQVTTVEEVTETV 485

(ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW,
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
PI WPI: 2003-02926/02.
XX DR N-PSDB; ACA42120.

XX New antisense nucleic acids, useful for identifying proteins or screening PT for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.
XX Claim 25; SEQ ID NO 66174; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of CC the 6213 antisense sequences given in the specification where expression CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid CC encoding a polypeptide whose expression is inhibited by the antisense CC nucleic acid; (2) a host cell containing the vector; (3) an isolated CC polypeptide or its fragment whose expression is inhibited by the CC antisense nucleic acid; (4) an antibody capable of specifically binding CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular CC proliferation or the activity of a gene in an operon required for CC proliferation; (7) identifying a compound that influences the activity of CC the gene product or that has an activity against a biological pathway CC required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological CC pathway in which a proliferation-required gene or its gene product lies CC or a gene on which the test compound that inhibits proliferation of an CC organism acts; (9) manufacturing an antibiotic; (10) profiling a CC compound's activity; (11) a culture comprising strains in which the gene CC product is overexpressed or underexpressed; (12) determining the extent CC to which each of the strains is present in a culture or collection of CC strains; or (13) identifying the target of a compound that inhibits the CC proliferation of an organism. The antisense nucleic acids are useful for CC identifying proteins or screening for homologous nucleic acids required CC for cellular proliferation to isolate candidate molecules for rational CC drug discovery programs, or for screening homologous nucleic acids CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*, CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of CC the target prokaryotic essential genes. Note: The sequence data for this CC patient did not form part of the printed specification, but was obtained CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

Query Match 16.1%; Score 386.5; DB 6; Length 682;
Best Local Similarity 27.5%; Pred. No. 1.9e-19;
Matches 95; Conservative 77; Mismatches 140; Indels 33; Gaps 4;
Qy 144 LLDAADDWVADRGEEAAAVDELVARFLPMKLTFDQQIAMDDYDSYAQRLHEDIDS 203
Db 356 LLDEIAD--LADGDTVAAATYE-----DFTGTAADSINYSDQ 392
Qy 204 ROELA---NAVATHVEAPLSSLEATSQDVAAERTDTMARTDDQVDRMADSREISSVSAS 260
Db 393 LRELVETINQTAQVQVAAAQETQSFMHLAE-----ASEHQAEQIAQGSAAINEMAVS 445
Qy 261 VEEVASTADDVVRTSSEDAEAIAQGCEAAADDALATMIDIDEATDGVTTAGVEQLGERAADV 320
Db 446 IDQVANASESSAVALRSVALANKNEVVNTITGNDNIREQIQTSKRKLGRSSQE 505
XX Sequence 682 AA;

Query Match 16.1%; Score 386.5; DB 6; Length 682;
Best Local Similarity 27.5%; Pred. No. 1.9e-19;
Matches 95; Conservative 77; Mismatches 140; Indels 33; Gaps 4;
Qy 321 ESVTGYYDDIAEQTQNMALANASIEAARAGEAGEFIAVVADEVKALAAEESREQSTIVEELY 380
Db 506 GDIVSLINDLADQTMNLNALNAAIQSMAGDAGRGGFIAVVADEVQRLAERSAAATKQIEALV 565
Qy 381 EQMQAEETEYTVDQDDEVNQRIGEYGRVEEAMETLQBTDAVEDAASGMQEVSTADEQA 440
Db 566 KTIQDTNEAVISMEQTTSEVVRGARLAQDGALEBIEKVKSTLAQIIONSNAAARQQA 625
Qy 441 VSTEVEVAEMVGDVDDRDGETAAALDDIADATDQQVTTVEEVTETV 485

Db	626 SSAGHSNTMVIQELTSOSAGTTATRS...GNLXKMASMRNSV 670	Db	264 ARAGEAKGPAVVADEBIRKLAESKKAAEDIRELIKQ-----IGDKIGESV 309
RESULT 6	AAB96483	Qy	406 ERVEEAMETIQEITDAVEDASGMVOSTATDBQVSTVEVAEM-----D 451
ID AAB96483	standard; protein; 435 AA.	Db	310 EVTQGAEVKTSIEVTKESVSYLTQVAEMMEVKA8LREKVYIQEGERKFILE 369
XX AC AAB96483;		Qy	452 GVDDRGEIAAALDDIADTQDQVRTEVEYRETGVKL 488
XX		Db	370 NLAASAEETAAEVSAAAEQSSALQELRESVRL 406
DT 29-Oct-2001 (first entry)		RESULT 7	
XX DE Putative sensory transduction histidine kinase and response regulator #2.		AAB96709	
XX KW Hyperthermophilic archaeon; hyperthermophilic protein.		ID AAB96709 standard; protein; 501 AA.	
XX OS Pyrococcus abyssi.		XX AC AAB96709;	
XX PN FR2792651-A1.		XX DT 29-OCT-2001 (first entry)	
XX PD 27-OCT-2000.		XX DE Putative sensory transduction histidine kinase & response regulator #4.	
XX PF 21-APR-1999; 99FR-00005034.		XX KW Hyperthermophilic archaeon; hyperthermophilic protein.	
XX PR 21-APR-1999; 99FR-00005034.		XX OS Pyrococcus abyssi.	
XX PA (CNRS) CNRS CENT NAT RECH SCI.		XX PN FR2792651-A1.	
PA (IFRE-) IFREMER INST FR RECH EXPL MER.		XX PD 27-OCT-2000.	
XX PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;		XX PF 21-APR-1999; 99FR-00005034.	
PI Querellou J, Weissenbach J, Saurin W, Heilig R;		XX PR 21-APR-1999; 99FR-00005034.	
XX DR WPI: 2001-126236/14.		XX PA (CNRS) CNRS CENT NAT RECH SCI.	
XX XX PA (CNRS) CNRS CENT NAT RECH SCI.		PA (IFRE-) IFREMER INST FR RECH EXPL MER.	
PT New nucleotide sequences isolated from Pyrococcus abyssi encode proteins		PI Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;	
PT useful in industry.		PI Querellou J, Weissenbach J, Saurin W, Heilig R;	
PS Claim 7; Page 1189-1190; 1657pp; French.		XX WPI: 2001-126236/14.	
XX XX PS Claim 7; Page 1189-1190; 1657pp; French.		XX DR WPI: 2001-126236/14.	
CC The present invention relates to the genomic sequence of Pyrococcus abyssi (see AAF86431 and AAH41223-7) and P. abyssi protein.		XX XX PT New nucleotide sequences isolated from Pyrococcus abyssi encode proteins	
CC a hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal vents. The present sequence is one such P. abyssi protein.		CC PT useful in industry.	
CC The proteins of the present invention have various potential industrial uses, since the proteins are stable at very high temperatures, some up to 110 degrees centigrade. Note: This patent is in the same patent family as WO2005062, which contains additional sequences as shown in AAB99132- CC WO2005062, which contains additional sequences as shown in AAB99132-		XX PS Claim 7; Page 1469-1471; 1657pp; French.	
CC Sequence 435 AA;		XX CC The present invention relates to the genomic sequence of Pyrococcus abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is a hyperthermophilic archaeon, which is isolated from deep sea hydrothermal vents. The present sequence is one such P. abyssi protein. The proteins of the present invention have various potential industrial uses, since the proteins are stable at very high temperatures, some up to 110 degrees centigrade. Note: This patent is in the same patent family as WO2005062, which contains additional sequences as shown in AAB99132- CC WO2005062, which contains additional sequences as shown in AAB99132-	
SQ 107 EYDTEYYAARARIGKIHDVGLGPPDVYLGMTRYTYGTLLDALADDVADRGEAAAVDE 166		XX CC Sequence 501 AA;	
Db 63 EYEKQLOREDEIYKILDRAQG_DL-----SVEKEIGMMGAKYRGQIED 107		Query Match 15.9%; Score 380; DB 4; Length 435;	
Qy 167 LVAREFLPMQLRLTFDQQIAMTDYIDSAYAQLRHEDIDSROELANAV-ATHEAPLSSLEPAT 225		Best Local Similarity 27.5%; Pred. No. 3.1e-19;	
Db 108 L-RKSISQLWNINNAIDDRNHTKVMKENIEQYADSQVAAEINQNSIEA----ORE 161		Matches 109; Conservative 91; Mismatches 129; Indels 12; Gaps 1.2;	
Qy 107 EYDTEYYAARARIGKIHDVGLGPPDVYLGMTRYTYGTLLDALADDVADRGEAAAVDE 166		Query Match 15.7%; Score 375; DB 4; Length 501;	
Db 63 EYEKQLOREDEIYKILDRAQG_DL-----SVEKEIGMMGAKYRGQIED 107		Best Local Similarity 24.4%; Pred. No. 8.6e-19;	
Qy 167 SQDVAAERTDPMRARTDDOVRMADSREIISVSVASVVEYTSTADVRTSDEAFALAQG 285		Matches 116; Conservative 111; Mismatches 175; Indels 74; Gaps 1.2;	
Db 162 QENISKMTDMR-----YHDISKE-TVSTMEFFRS---MRENAOLAKEGGKG 207		Qy 65 VTDIFYDHL-----SYERTQDLFANSTKTVEQKETQAEYLILGGRGYDTEAA 114	
Qy 286 EAAADDALATMMDIBATDQVTAQVEQLGRERADEVSTVQDIDIAEQTMNLAINATEA 345		Db 30 ISDIDNIEKTMKVVOQGSVEATRETLENAK---SOLKET----LWLSIGMSI---- 77	
Db 208 RQAADQ-IEISRMMEKIBETTVGAEAGK--SIENTINVISSIEQTNLAINATEA 263		Qy 115 QRARIGKIHDVGLGPPDVYLGMTRYTYGTLLDALADDVADRGEAAAVDELVAPP 174	
Qy 346 ARAGEAGEGFPAVVADEVKALAEESREQSTREEVELVEQMQAETEETVQDLEVNQRIGEV 405		Db 78 ---VAVISGALGN---RLMNSTMPINEMKIAESAEGLKSLRADMVKSQYRDEDE 129	
Db 130 IGKJIBGFRASQVLTLEITERM-EKISEGDDVSDDLKVHAKGDFESILNSMRKRIS 187		Qy 175 L-KULTFDQIAMDY--IDSYAQLRHLDEIDSROELANAVTH---EAPLSLEATSQ 227	

the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

X Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
X plant biocontrol agent.
X
X Acinetobacter baumannii.
X
X US6562958-B1.
X 13-MAY-2003.
X 04-JUN-1999; 99US-00328352.
X R 09-JUN-1998; 98US-0088701P.
X (GENO-) GENOME THERAPEUTICS CORP.
X
I Bretton G, Bush D;
X X WPI; 2003-576092/54.
R N-PPSB; ADA29759.
X New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
X for diagnosing a bacterial disease, as components of antibacterial
X vaccines, as targets for antibacterial drugs, or as biocontrol agents for
X plants.
X
S Example; SEQ ID NO 5172; 328pp; English.
C The invention relates to isolated *Acinetobacter baumannii* nucleic acids.
C The *A. baumannii* nucleic acids and polypeptides are useful as reagents
C for diagnosing a bacterial disease, as components of antibacterial
C vaccines, as targets for antibacterial drugs, to detect the presence of
C *A. baumannii* and other *Acinetobacter* species in a sample, in screening
C compounds for the ability to interfere with the *A. baumannii* life cycle
C or to inhibit *A. baumannii* infection, and as biocontrol agents for
C *baumannii* plants. The present sequence represents the amino acid sequence of an *A.*

Query Match	15.6%	Score	373.5;	DB	6;	Lengt	709;
Best Local Similarity	28.7%	Pred. No.	1.7e-18;				
Matches 100;	Conservative	Ni-smatches	141;	Indels	41;	Gaps	5;
Matches 100;	Y						
144	LLDAADDVYADRGEAAAAYDELVALEPMLKLITEDQIAMIIDYIDSYAQRIHDEIDS	203					
388	LLDDEAD--LADGDRSYATSE-	-	-	-	-	-	DFTGAIADSINFADQ
204	RQELANAVATHVEAPLSSLEATSQVAERTDTMRARTDDQVDRMADVSEISSVSASVVE	263					
425	YRDL-----VSRHETTSQFVARYTQDTSQNTLQAESPHQIASTARNIE	473					
264	VASTADDV-RRTSEDAEA-----LAQOGEAAADALATMTDIDEATDGUTAGYOLGER	316					
474	MAQSDFQSVSANASEAEVAQRVSQIASNGAQVNRSISIGMDTRIREQIOTESTKIRLQBS	533					
317	BADVEVTGVDDIABQTNMLAINASIEAARAGEFGEFAVVADEKVALAESREQSTRV	376					
534	SQEONIVSPLINDADQTNILALNVAIQIAQSMAGRGPFVADEVQRLAERSASATQI	593					
377	EELVEMQAEETEFTVDLDEVNQRIGEGVERVEAMTELQEITDAVEDAASGNQEVSTAT	436					
594	ETLVQTCUTNEAVLISMECTTTTVEVRGANLAKGIADEIQVSGDLAKLASICDAA	653					
437	DEQAVSTEEVAMYGVDDDRAGEIAAALDDIADATDOQVRVVERETV	485					
654	KLQSSASHIATTTWVQEQITSQTATDFTARSVSELANMAESLREVS	702					

RESULT 10
BU040238 D ABU040238 standard; protein; 686 AA.

AC ABU40238;
 XX DT 19-JUN-2003 (first entry)
 XX DE Protein encoded by Prokaryotic essential gene #25765.
 XX DE Protein encoded by Prokaryotic essential gene; cell proliferation; drug design.
 XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX KW Pseudomonas putida.
 OS XX
 PN XX
 WO200277183-A2.
 PD XX
 03-OCT-2002.
 XX PF 2002WO-US009107.
 XX PR 21-MAR-2001; 2001US-00815242.
 PR 21-MAR-2001; 2001US-00918993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW,
 Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA,
 XX DR N-PSDB; ACA44108.
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.
 XX PS Claim 25; SEQ ID NO 68162; 1766pp; English.
 XX CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or (9) a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp://wipo.int/pub/published/pct/sequences

Query Match 15.5%;
Best Local Similarity 24.7%;
Score 371.5; DB 6;
Prod. No. 2.3e-18; Length 686;

Qy	3.7	L SFTGIDD-DTMALAAEOPLFERATADALYTFDHDLESYERTODLF--ANSTKTEVQLK 93
Db	237	I QTVREVEDADARARLAEIAELFQFVAGSVD-----EILETSPPLFRYREAAGNFSIS 289
Qy	94	E TOAQEVLLGIGRGEVITEAAQRARIGKTHDVLGPDVYLGATRYTGLLDALADDYV 153
Db	290	Q TLDEASHLANG-----FENLAGRTLDTVG-----GYVGLL-ALASITL 33.0
Qy	154	- - - - - ADRGEAAAAYDELVARFLPMLKLTDPOQIA-----MD 187
Db	331	I GLGMVYRTTNROLRETAEKNERNOQIA-----MRUJDEBELGDLTIVVSYTE 383
Qy	188	T YIDSQAQRLLHEDIDSQRELA--NAVATHVEAPLSSLEATSQDVAERTDTMRARTDDQV 24
Db	381	DFTGATADSINTSYSDQQLIRDLVATINNSAVQOAAVQDTONTAROLAKASEHQAOQISEAS 44
Qy	245	DRMADVSEREISSVASVEEVASTADVRRTSEDAEAALAQOSEAAADDALATMTDIDEATD 30
Db	441	E AVGDMWESIDRVSAHAYESAKVA-----ERSVIAIANGNEVHTNTNGMDNIREQIQ 49
Qy	305	GVTAGYEQLLGRERADEVSYTGVDIIDAETQTMALANASITAARAAGEGFAYVADEKVA 364
Db	494	D DTAKRIRKLGESSQBEGDVISSLDDADQTNLANAAQSLAEGBAGRFAYVADEQVR 55
Qy	365	LAEFRESREQUESTREVELYEQMQAETEETYDQDDEVNORIGEYVERVEAMETLQEETDAVED 424
Db	554	LAERSSSSATRQLEALVRTICDNEAVISMEOQTAFUVGARLAQDAGVALAEFGVSON 61
Qy	425	A ASGMQEWSATADEQAVSTEEVAEMYVGVDDRAGEIAAALDIADATDQVRTVEEVR 48
Db	614	L ADLHLSISDAAQLOQTSSAGQTSHTMVAIQTTAQTSAGSGATADSIRHLARMASEMRRS 672
Qy	485	V 485
Db	674	V 674

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expressed by the nucleic acid inhibiting proliferation of a cell. Also included is (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically bind the polypeptide; (5) producing the polypeptide; (6) inhibiting cell proliferation or the activity of a gene in an organism required for proliferation; (7) identifying a compound that influences the activation of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product 1 or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by the target prokaryotic essential genes. Note: The sequence data for patent did not form part of the printed specification, but was obtainable in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

Query Match	Loc.	Score	Loc.	Score	Length	Qry
Best Local Similarity	22.7 %	Pred.	No. 1. e-17;			
Matches	120; Conservative	98; Mismatches	212;	Indels	98; Gaps	14;
Qy	28	LDEAEIARLISFTGFLDDDTMAAL---	AAEOPLF-		57	
Db	139	LNQODGWPGFPIGSKPTVSMFQDANGEYKMFAMYQLVNQNTMSMMSKSMDDMWRLIN	198			
Qy	58	---EATADALVTDFYDHLESVERTQDLPFANSTKTVEQLKETQAEYVLLG-LGRGEYDTEY	112			
Db	199	SFKIEDTGFVFLNQGEVQHRCQV	--KSSLQQIYGSASALLNKSGNLNISTDY	254		
Qy	113	AAQRARIGKHDVLGLGPDPDLYGAATRYVTCOLL--DALAD-DVNAIR--GEEAANAVD	165			
Db	255	QGEEMWASI-----YIESMMWFLVSTVPEVFAELDAVAGRMMLTTLVAAIF	304			
Qy	166	ELVARFLPMLKLITPDQOIAMDTYIDSAYAQRLHD-----BIDSROELA----	208			
Db	305	IFMGFLA-----NSIAMP--INQIARKEFTDLERGDGLSQRREVKGNDENIAQLSKG	354			
Qy	209	-NAVATHVEAPLSSLEATSQD-----VAERTDTMRARTDDQVDRMADVSEISSVSYS	260			
Db	355	FNGFEKIKQSOKDQAHTNSOCQDRTQIVWTAINQNGAT	414			
Qy	261	VEEVISTADUVRRTSDEAANLAQCGERAADDALATMTIDIDATGTVGQELGERAADU	320			
Db	415	ISELASNAATAAETANQASGRADQGENVNKAESLRALHDIENTSKVVEQLASTTQEII	474			
Qy	321	ESVTSTGIDDIAEQTNALNASEEARAGEAGEFIAVDAEYKALAESREBOSTREVELV	380			
Db	475	GSILIDIRGSEQINTLNALNATEARAGDORGFAVDAEYRNLAERTASSTEIQM	534			
Qy	381	EQMQAETEETYDQDLEVNORIEGEYVERVEAMETQEIITDAEASGCMQEVSSTATDEQA	440			

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway; (8)
 CC required for proliferation, or that inhibits cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the best compound that inhibits proliferation of an
 CC compound; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp://wipo.int/pub/published_pct_sequences

XX SQ Sequence 539 AA;

Query Match 14.9%; Score 356.5; DB 6; Length 539;
 Best Local Similarity 26.6%; Pred. No. 2.1e-17;
 Matches 126; Conservative 81; Mismatches 153; Indels 113; Gaps 15

Qy 38 SPTGIDDDTMAL----AAEPLPLRATADLTVTDPYDHLSEYRTQDLFANSTKTEQL 92
 Db 126 SFAASIDEMMPLAERLNDDAYLQLKNUVKAGQSAAAPAAARLGEFSTNDKESSETLAAHET 185

Qy 93 KETOAYYLGLGRGEYDTEYAAQRARIKGTHDVLGLGPDPVYLGATRYTYGFLDALADDV 152
 Db 186 RETMIVYV-----YAA-----LLLIVGVTAGSLYMTAVRPLQR-- 222

Qy 153 VADRGEEAAAANDELYARELPMLKJJTEDQIAMDITYDSYAQRJHEDISROELANAYA 212
 Db 223 -ASQOFERMAAGD-LSARI-----EGNSNEIGAILT 252

Qy 213 T--HVEAPL-----SLEATSDQVAERTDMARTDDQVDEMADVSREISSVSA 259
 Db 253 ALRHMQESLPTVAAVRGVDEINCSREIAGNTTSSRPEQ-----AASLERTAA 305

Qy 260 SVEEVAST---ADDVTRTE --DAEALAOQGEAAADDALATMDIDEATDGVTAGEQ 312
 Db 306 SMEQLASTVKONADNRQANOLAAASDVAE5GGSAVEVSTM-----DGISASRR 358

Qy 313 LGERAADVESVTGVIDIAEQTMNLMNSTEARAGEGFAVVADEVKALEESREQ 372
 Db 359 ISE-----IVSVLICIAFOTNILALNRAVEAARRAGEQGKGFAVAGEVRSIAQRQA 411

Qy 373 STRVVEELBQNQAEETFDQLDENVORIGEGVERVEAMETLOQITDAVEDAASGMQE 432
 Db 412 AKEIKVUIE-----DSVD----KVGTTSQKVFRAGAMQE YAVSVKVTDTNGEI 457

Qy 433 STATDEQAVSTEEVAAEMVGVDYDRAGHIAALDDIADATDQQVTRYVEBREVY 485
 Db 458 SAASBEOSSGIEQVNRAVSQMDETOQNAALVVEAAAGAAGLSQQAQLRAAV 510

RESULT 15
 ABU49142
 ID ABU49142 standard; protein; 626 AA.
 XX AC ABU49142;
 XX DT 19-JUN-2003 (first entry)
 XX CC Protein encoded by Prokaryotic essential gene #8628.
 XX CC Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX CC Bordetella pertussis.
 WO20020277183-A2.
 03-OCT-2002.
 21-MAR-2002; 2002WO-US009107.
 XX CC (ELIT-) ELITRA PHARM INC.
 PR 21-MAR-2001; 2001US-00815242.
 PPR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-034923P.
 PDR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 WPI; 2003-02926/02.
 N-PSDB; ACA26971.

Claim 25; SEQ ID NO 51025; 176pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway; (8) required for proliferation, or that inhibits cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the best compound that inhibits proliferation of an compound acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

RESULT 15
 ABU49142
 ID ABU49142 standard; protein; 626 AA.
 XX AC ABU49142;
 XX DT 19-JUN-2003 (first entry)

XX	Protein encoded by prokaryotic essential gene #34669.	Db	298	MSPLKTLDSAIKDIASSGGDLTKLKDNTLDKEFSELALGENSETMELGSQIROLKTIASG 357
DE		Qy	222	LEATSDQVDAERTDTMRARTDDQVDMADSVREISSVASVEEVASTADDYRRTSDEAEL 281
KW	prokaryotic essential gene; cell proliferation; drug design.	Db	358	VLDGAEKTANEAVSRIVVEQQL----QELSQLATMNEMAMTASEVANSAQVADA 410
OS		Qy	282	AQGEEAAADDALATMDIDEATDGVT-----AGVEOLGE--RAAD-YESVTGVDIDIA 331
XX	Vibrio cholerae.	Db	411	AREGESA--SLEGSVVHETTDALQRUSIRIGSVEDYKELVATDRIETVLDVINDIA 467
PN	WO200277183-A2.	Qy	332	EQTNMLALNNSIEARAGAGEGFAVVADEVKALAESESREQSTRVLELYQMQRETEETY 391
XX	PD 03-OCT-2002.	Db	468	DQTMFLALNVAIEARAGSGRGFAVVADEVRVLAQRTQSTMQSEILLEQLQEGAKNVS 527
PF	21-MAR-2002; 2002W0-US009107.	Qy	392	DQLEDEVNQRIGEGERVEEAMETLQEITDAVEDIASGMDEVSTATDEQAVSTEEAEMD 451
XX	PR 21-MAR-2001; 2001US-00815242	Db	528	RSMDESKLETDIVVEKTNYNEKISLVLQQQATHRTSDMNLQIASAEQSLVABBINNTV 587
PR	06-SEP-2001; 2001US-00948993	Qy	452	GVDDRAEIAAALDDIADATDQQVTRVERVRETVGK 487
PR	25-OCT-2001; 2001US-0342923P	Db	588	NIKOLSIKUSENASAQTMENAQVSKVREQNLINE 623
PR	08-FEB-2002; 2002US-00072851			
PR	06-MAR-2002; 2002US-0362699P			
XX	PA (ELIT-) ELITRA PHARM INC.			
XX	PT Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;			
PI	PT Wall D, Trawick JD, Carr GD, Yamamoto R, Forysth RA, Xu HH;			
XX	DR WPI: 2003-029926/02.			
XX	PS DR N-PDSB; ACA53012.			
XX	PT New antisense nucleic acids, useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs.			
XX	PT			
PT				
PT				
PS	Claim 25; SEQ ID NO 77066; 1766pp; English.			
XX	The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular proliferation or the activity of a gene which the test compound that influences the activity of the gene product or that has an activity against a biological pathway (7) identifying a compound that inhibits proliferation of an isolated nucleic acid; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than <i>S. aureus</i> , <i>S. typhimurium</i> , <i>K. pneumoniae</i> or <i>P. aeruginosa</i> . The present sequence is encoded by one of the target prokaryotic essential genes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences			
XX	Sequence 626 AA;			
SQ	Query Match 14.5%; Score 348; DB 6; Length 626; Best Local Similarity 27.4%; Pred. No. 1e-16; Matches 92; Conservative 78; Mismatches 134; Indels 32; Gaps 6;			
Qy	174 MLKLLTFDQQI-----AMDTYIDSAYQLHDEDSQELANAVATHVEAPLSS 221			

B
I
A
N
T

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 10, 2004, 15:23:41 ; Search time 24.8644 Seconds
1015.311 Million cell updates/sec (without alignments)

Title: US-09-455-978B-2

Perfect score: 2394

Sequence: 1 MSNDNNTLVADVRNGIDGH.....ATDQQVRTVEEVRETVGKLS 489

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 10%

Listing first 45 summaries

Database : Issued Patents AA:
 1: /cgn2_6/ptodata/2/iaa/5a_COMB.pep.*
 2: /cgn2_6/ptodata/2/iaa/5b_COMB.pep.*
 3: /cgn2_6/ptodata/2/iaa/5a_COMB.pep.*
 4: /cgn2_6/ptodata/2/iaa/65_COMB.pep.*
 5: /cgn2_6/ptodata/2/iaa/5ctus_COMB.pep.*
 6: /cgn2_6/ptodata/2/iaa/backfiles/pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	412.5	17.2	640	4	US-09-252-991A-23252	Sequence 23252, A
2	401.5	16.8	545	4	US-09-252-991A-31086	Sequence 31086, A
3	398	16.6	663	4	US-09-252-991A-23255	Sequence 23255, A
4	397.5	16.6	906	4	US-09-252-991A-32175	Sequence 32175, A
5	389.5	16.3	696	4	US-09-252-991A-16965	Sequence 16965, A
6	380.5	16.1	710	4	US-09-252-991A-32789	Sequence 32789, A
7	380.5	15.9	734	4	US-09-252-991A-30703	Sequence 30703, A
8	380.5	15.9	857	4	US-09-252-991A-23956	Sequence 23956, A
9	379	15.8	701	4	US-09-252-991A-23288	Sequence 23288, A
10	373.5	15.6	709	4	US-09-252-991A-5172	Sequence 5172, A
11	373	15.6	614	4	US-09-252-991A-31412	Sequence 31412, A
12	368.5	15.4	573	4	US-09-252-991A-18744	Sequence 18744, A
13	366	15.3	760	4	US-09-252-991A-31724	Sequence 31724, A
14	361	15.1	613	4	US-09-252-991A-26180	Sequence 26180, A
15	359	15.0	579	4	US-09-252-991A-18264	Sequence 18264, A
16	358.5	15.0	611	4	US-09-252-991A-20097	Sequence 20097, A
17	352.5	14.7	563	4	US-09-252-991A-31048	Sequence 31048, A
18	348.5	14.6	684	4	US-09-252-991A-28604	Sequence 28604, A
19	346.5	14.5	504	4	US-09-252-991A-26180	Sequence 26180, A
20	343.5	14.3	537	4	US-09-252-991A-20929	Sequence 20929, A
21	334.5	14.0	653	4	US-09-252-991A-18264	Sequence 18264, A
22	334	14.0	523	4	US-09-252-991A-8138	Sequence 8138, A
23	333	13.9	645	4	US-09-252-991A-16799	Sequence 16799, A
24	329	13.8	548	4	US-09-252-991A-23147	Sequence 23147, A
25	321.5	13.4	535	4	US-09-252-991A-4593	Sequence 4593, A
26	320.5	13.4	670	4	US-09-252-991A-26867	Sequence 26867, A
27	315	13.2	595	4	US-09-252-991A-6908	Sequence 6908, A

ALIGNMENTS

RESULT 1
 US-09-252-991A-23252
 ; Sequence 23252, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; FILE REFERENCE: 107196.136
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; SEQ ID NO: 23252
 ; LENGTH: 640
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-23252

Query Match 17.2%; Score 412.5; DB 4; Length 640;
 Best Local Similarity 24.4%; Pred. No. 3.3e-24;
 Matches 149; Conservative 97; Mismatches 199; Indels 165; Gaps 19;

Qy 7 TLVTPADVRNGIDG---HALADRIGL---DEAEIAWRLSFTGIDDITMAALAAEQLIF 57
 Db 59 TLTAASN1QSWLEGRMHLVEGLASQLLADQPDRAANIAQL-----EQPVF 103

Qy 58 -----EATA DALY TD PY DHL-ESE Y-RTO DLFANS-----TKTVEQ 91
 Db 104 SRNFASVYLGEEASGTFTMRPYDAMPEGYDPRTRAWYKDALAADRLLTYPFPYDAGTGHQ 163

Qy 92 -----LKETQAENYLGLGRGEYDTE-----YAQRARIGKTI-HDV LGL-128
 Db 164 ILAMS LPLYR HAG QLGV AFG AYFL VSD GKG KILLHPDSGLV 223

Qy 129 -----GPDVYLGAY-----TRYTGILDD-----TRYTGILDD-----146
 Db 224 LKTLAENAPKGANIVPGYHEVLDGGSQFSVTPVKGTVWVAVLVDRTAYSMIL 283

Qy 147 -----AlADDVVA-----DRG-----EEAAA AVDELVARFLPMKLL 178
 Db 284 EFRTSAIVATLIAVVGMLLGMMLIRYLMQPLTDGMQDLAQGECSDLTKR---LKVT 339

Qy 179 TFDQOQIMDTYLDSYAQLHEDLSRQBLANAVATHYEAPLSSLEATSQDV AERTDMRA 238
 Db 340 SNEDEFGTLANAFNRVERIHEST--REVAGTA----RQLHDV AQLVVNASNSMSA 388

239 RTDDQVDRMADVSRRISSVSASVEREASTADDVVRTSEDAEAALAQGEAAADDALATMTD 298
 389 NSDEQSNRNTSVAIAINELGAAQIARNADASHASNDHQADGKQYVEQTIRAME 448
 299 IDEATDGVTAGVQGLERAADVESVTGVIDDIAETQNMALNASTEARAAGEAGEGFAYV 358
 449 LSEKTSASCANIEALNSRTVNLNIGOLLEVKGSEQTQNLLALNAAEARAAGEAERGFAYV 508
 359 ADEVKALESREQRSTRVLELVQMQAEETVTDQDLEVNQIGGVKEEAMETLQEI 418
 509 ADEVNLRAHQAEQSQIQMIEELQIGAQAEAVSMTEORYSLESVETANRAGERLGSV 568

Db 375 IHSIIQVKIGIAEQUNLLALNAAEARAAGEAGEGFAYVDAERTQSTOEITAM 434
 Db 380 VEQMGAETETTVDQDLEVNQIGESEVEREEAMETLQEITDAEDAAQGMEQEVSTATDEQ 439
 Qy 435 TERTFASTGQAISMEAGYSRVNEGVFSARQAGYSINETLDGTRHAASTYDEDISCTIREQ 494
 Db 440 AVSTBEAENVGVDGVIDDRAGEIAAALDDIADATDQOVRTVEBRETIV 485
 Qy 495 SRASDEIAQRVELAQRSQNTQAHEMA AT--ARRINEVAATM 536

Db RESULT 3
 US-09-252-991A-23255
 ; Sequence 23255, Application US/09252991A
 ; GENERAL INFORMATION:
 ; PATENT NO. 6551795
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.116
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIORITY NUMBER: US 60/074,788
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIORITY NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 23255
 ; LENGTH: 663
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 28.3%; Pred. No. 4.7e-23;
 Matches 118; Conservative 74; Mismatches 171; Indels 54; Gaps 10;

Db 103 LGRGLEYDTEAAQRARIKGTHDVLGL-GPDVYLG-----AY---TRYTGGLLADADD 151
 ; SEQ ID NO 23255
 ; LENGTH: 663
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 28.3%; Pred. No. 4.7e-23;
 Matches 118; Conservative 74; Mismatches 171; Indels 54; Gaps 10;

Db 261 IGSFFSEALHG-NTRILSPSPVKGLSLDWYIGSVDRDKAYANLTKLRTSAIIYALIA 319
 ; SEQ ID NO 23255
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 192 SYAQRLHDEIDSROQLANAVATHVAPLSSLEATSDVAERTDTMARTDQVDMADVS 251
 ; SEQ ID NO 23255
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 376 RFVERTHESI---REVAGTA-----ROLHDVAOLVNVANSNSNSDFQSNTNTNSVA 424
 ; SEQ ID NO 23255
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 252 REISSVSYSEVEVASTADDVVRTSEDAEAALAQGEAAADDALATDIDEATDGTTAGVE 311
 ; SEQ ID NO 23255
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 425 AAINELGAAQETIARNAADASHASDHQAEDGKVQEVTIRANNELESKISSCANIE 484
 ; SEQ ID NO 23255
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 312 QLGERAADVDESVTGVDDIAEQTMNLNALNAAEARAGEAGEGFAYVVADEYKALARES 371
 ; SEQ ID NO 23255
 ; LENGTH: 545
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 485 ALNSRTVNGQILEVKIGSEQTNLLNALNAAEARAGEAERGFAYVVADEYKALRAQE 544
 ; SEQ ID NO 23255
 ; LENGTH: 544
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 372 QSTRVBEELVQMQATEETVTDQDLETVNORIGEGVVEREEAMETLQETTDAEDAAQGMQE 431
 ; SEQ ID NO 23255
 ; LENGTH: 544
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 545 SAQQIQMIEFLQIGAQQEAVSTMTMQRYSLESVETANRAGERLSSVTGRIAEDDMNQS 604
 ; SEQ ID NO 23255
 ; LENGTH: 604
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 432 VSTATDEQAVSTEEVAVEMVGUDDRGEIAAALDDIATDQOVRTVEEVRVYKL 488
 ; SEQ ID NO 23255
 ; LENGTH: 488
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 605 VATATERQTA-----VVDSLNMIDTEINTLNQEVENIQTATLRAQGELETQASRGL 654
 ; SEQ ID NO 23255
 ; LENGTH: 654
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 315 NISQISDHADNARVISAKSEELESSCCGIVLNVEGMSRIADVNQSSSTSITLGQSSDE 374
 ; SEQ ID NO 23255
 ; LENGTH: 374
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

Query Match 16.6%; Score 398; DB 4; Length 663;
 Best Local Similarity 24.9%; Pred. No. 1.9e-23; Gaps 16;
 Matches 131; Conservative 111; Mismatches 205; Indels 79; Gaps 16;

Db 320 VESVTGYDIAEQTMNLNALNAAEARAGEAGEGFAYVVADEYKALAESESREQSTVEEL 379
 ; SEQ ID NO 23255
 ; LENGTH: 379
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-23255

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 32715
; LENGTH: 906
; TYPE: PRT
; ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-16965

LENGTH: 696
; TYPE: PRT
; ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-16965

Query Match 16.3%; Score 389.5; DB 4; Length 696;
Best Local Similarity 25.3%; Pred. No. 2.3e-22;
Matches 131; Conservative 98; Mismatches 210; Indels 79; Gaps 14;

Qy 1 MSNDNTLVTADVRNGIDGHALADR1--GLDEAETIAWRLSFTGIDDITMAALAAEQPLFEE 58
Db 213 IAGDENSVOAQAND-SFRGDA-SLFGRVILKGNOEGNAMSISKV-TNAEAVDRNLNEAELFEE 269

Qy 59 ATADAYTDFYDHLSSYERTODLF --- ANSTKTYBOLKETOQAXYL--- GLGREYD 109
Db 270 ---- FVSGSYD-B-EILLETSPLDFYREANNIFSVQSTLKDASQLADGFENLAGGRSI 322
Db 323 NLFA----GYVLGALALASILIGLVMTRETNRLAETAEKNDRNQAALRLDEIAD 376

Qy 110 TEYAAQBARIGKHDVGLGPDVYLYATYYTG -- LIDALAD 150
Db 377 -LAQDLTVATVE-----DFGAIADSINSDQRELVET 413

Qy 151 DVAADGEAAAAYDDELVARFLPMIKLLTFDQIQAMDTYIDSYAQRHLHEDIDSROBLA-- 208
Db 209 -NAVATHVEAPLSSLEATSQDVAERTDTMRARTDDQVDRMDSREISSVSASVEEVAST 267

Qy 414 INQTAQVAAAQETQSTAMHAE -----ASEHQAEQTAGASAANENAVSIDQVSAN 466

Qy 268 ADDVRRTSDEAALAAQGEAAADDALATMDIDEATDGVTAGVEQGLGERAADVESVTGVI 327
Db 467 ASESSAAERSVATANKGEVNHNTITGMNDNIREQOIDSKRKGESSOEIGDIVSLI 526

Qy 527 NDIAQTNLNALNAIQASMAQDAGRFAVDEVQLAERSSAATKQBALVKQTDT 586

Qy 388 EETVDQDLYNQRIQGEQVERVEAMETLQEITDAVEDAASGMQEVSTADEQAYSTEVA 447

Qy 328 DDIAEQTNMALNAAARAGEAGEFAYVADEKVALAESRECOSTRVEELVROMQAET 387

Db 528 NEAVISMEQTSEVRGARIAQDAGVSKTLALIQNISNAAQRQQASSAGHIS 646

RESULT 6
US-09-252-991A-32789
; Sequence 32789, Application US/09252991A
; PATENT NO. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEARIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIORITY NUMBER: 1999-02-18
; PRIORITY NUMBER: US 60/074,788
; PRIORITY NUMBER: 1998-02-18
; PRIORITY NUMBER: US 60/094,190
; PRIORITY NUMBER: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 32789
; LENGTH: 710
; TYPE: PRT
; ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-32789

Query Match 16.1%; Score 385; DB 4; Length 710;
Best Local Similarity 27.8%; Pred. No. 5.e-22;
Matches 131; Conservative 83; Mismatches 211; Indels 42; Gaps 10;

Qy 39 FTGIDDDITMAALAAEQPLFEAT-----ADALVTDFYDHLSEYERTQDLFANS-- 85

RESULT 5
US-09-252-991A-16965
; Sequence 16965, Application US/09252991A
; PATENT NO. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEARIC ACID AND AMINO ACID SEQUENCES RELATING TO *PSEUDOMONAS*
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIORITY NUMBER: US 60/074,788
; PRIORITY NUMBER: 1998-02-18
; PRIORITY NUMBER: US 60/094,190
; PRIORITY NUMBER: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 16965

Db 257 FERSDKAQAAFAADALRQAATTLRGQLPGEADALEQAMGSLGQYRRGGIEQFRAGVIR 316 Qy 254 ISSVASVERVASTADDVRR-----TSEDAEAIAQCGERAADDALATMDIDBATDGV 306 Db 491 IELVATPAHENTATAQDVARNTHAEAAAHDAQAHQKQIVESSAAQALASEIGRA 550 Qy 86 TKTVEOLKETOAEYLGLGR----GEYDTYYAAQRARIGK-IHDVLGLGPBDVTLG-AVT 138 Db 307 TAGVBOLGERAADVESVTGVYDDIAEQTNMLALANASIEAARAGEEGFAYVADVKALA 366 Qy 3117 TRQAQQAMQS-TQDMARA9GRILTEGRQLRESTASDRASIWLIALAFLGCVGWAIN 376 Db 551 VGVVQNLAKDSENTINAILVARGIAEQTNLLANAAIEAARAGEGQRFQAVADEVRNLA 610 Qy 139 RYTYGLD-ALADDVVAADRGEAAAADVDELVARFLPMKLKLTEDQIAMIPTYIDSYAQL 197 Db 367 BESSROSTRVEELVQMOQATEETVTDQLEVNQRGEGVERVEEAMETLOBTTDAVEDAA 426 Db 377 RQIVRPEDEALA-----DAAEAAGDILGKRPNPLTQRDEL-----GQLQRV 420 Qy 611 QKTTQATEEIQSMICIQLQQCTRDVVKVMQDSQERTDSVRHAROAEEALESITQAVSVIN 670 Db 421 MORMGDSLRELVGRIGDGV----SOLASSABELSAVTEOTRAGVNSQKVTDOVATMHE 476 Qy 427 SGMOEVSTAVDQAVSTEEVAEMVGSVDDAGEIIAALDDIATDQQTVEVERETGV 486 Db 477 MAATVQDVARNELASQARQADEEARQGDAVDQAVTRIERLASEMDVSSRAMARKNE 536 Qy 671 DMNPQITASAAEQSQAVIDASQRAESQAGADEASQRAESQRLVYN 730 Db 487 K 487 Db 731 Q 731

RESULT 8
US-09-252-991A-23956
; Sequence 23956, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 23956

Query Match 15.9%; Score 380.5; DB 4; Length 857;
Best Local Similarity 26.8%; Pred. No. 1.6e-21;
Matches 141; Conservative 92; Mismatches 226; Indels 67; Gaps 16;

Qy 2 SNDNDTLYTADVRNGIDGHALADRIGHD-----EAEIAN-RLSFTGIDD-----DTM 47
Db 352 ANSLDIQSSLAEDG-A-YAUKLKSGHIDGEPLQQAQVAQVNRLVQGLDEARSRLEAR 409

Qy 48 AALAAQP--LFEA-----TADALYTDFYDHLESYERTQDFNSTKTVQELQK 93
Db 410 AADAQEGCPKSHREALELAARLEQATTDAYSVWVTDLNTNRGFDPLAE-YRASSQLQ 467

Qy 94 ETQAEYLGLGRGE---YDTEYAAQBARIGKIHDLGLGPDVYLGATRYTGYLDA 149
Db 468 ERQMYAAMGERAGQVMARVDRSWEQQQM-LHS-LRTNSLIVGR-----AVULLV 518

Qy 150 DVVADRGEEAAAVADLVARFLPMKLKLTEDQOQIA--MDTYIDSYAQRULHDEDSRQE 206
Db 519 -----GLGAAPFGISLIVR-PLRQAMGYAHRIAEQDGLAVRVDs--ERRDEVGOLMA 566

Qy 207 LANAVATHVEAPLISLE-----ATSODVAERTDTMARTDDQVDRMADSREISSVSA 259
Db 567 AVRMATGSLRGTVSLODGVRIGASEALSQVTRFLGDSQRATEQVATAMQMAA 626

Qy 260 SVEEVASTADYVRRTSDEAIAQCGERAADDALATMDIDBATDGVTAGVEOLGERAAD 319
Db 627 TVHEVALNAEEAAGAAEADGKVSYCQBVVRTLERLERAFAVRAATASVALSASVRL 686

Qy 320 VESVTGVYDDIAEQTNMLALANASIEAARAGEEGFAYVADVKALAESREQSTRVEL 379

RESULT 7
US-09-252-991A-30703
; Sequence 30703, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 30703
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-30703

Query Match 15.9%; Score 380.5; DB 4; Length 734;
Best Local Similarity 24.7%; Pred. No. 1.3e-21;
Matches 119; Conservative 91; Mismatches 178; Indels 93; Gaps 13;

Qy 41 GIDDITMVAALAAEQLFFPATADALVTDYDHLESY-----ERTQDFPANSKTKTE 90
Db 310 GNEVDNLKNTLVTDQPLYIDAE-----HGHIELFLPFITIADSGVRWTMLQIPQAAVFG 363
Qy 91 QLKETQAFY-----LIGLGRGEYDTHYYAAQRARIGKIHDLGLGPDVYLGAY----T 138
Db 364 BIQQLQGELSQRQQDINGM-----SLAGUVVAALGL-LVWVLGYGIARPL 409
Qy 139 RYTYGLDIALAD---DVAADRGRGEAAAADVDEL-VARFLPMKLKLTEDQOQIAMDITYDSY 193
Db 410 RQFLGMLDIAQCGEGLTRRLSSRA-----DELGIAK-----GFNTFUGKL 452
Qy 194 AQLHDELDSDQELANAYATHVEAPLSSLEATSDODVAERTDTMARTDDQVDRMADSRE 253
Db 453 QNNIGQVQSVOKVS-----DSSHTADIAIRTNQGVQQQL--AE 490
Qy

Db	687	IGSVLVDVTKSVAEQTNLLALNAAEAARRGGDQGRGSEFAVDAEVRALARTRQQSTAEIELU	746	; CURRENT FILING DATE: 1999-06-04 ; NUMBER OF SEQ ID NOS: 8232 ; SEQ ID NO: 5172 ; LENGTH: 709 ; TYPE: PRT ; ORGANISM: Acinetobacter baumannii US-09-328-352-5172
Qy	380	VEQNQAEETETVDOLEDEVNQRIGEVVERVEAMETLQLETITDAVEDDAASGMQEVTATEDEQ	439	
Db	747	IGALONGTQCAVQRMORSHOLVQSVDDAQLTEALGNITATAVLIQMNQIAASEQQ	806	
Qy	440	AVSTBEVAEMVGDYDRAEGFIAAALDDIADATDQVTRVEVRETY	485	
Db	807	SAVABEINRSVTAIREVAQDQAQMNSTASSSEQLAELGRRELQMV	852	
RESULT 9				
;	Sequence 23288, Application US/09252991A			
;	Patent No. 6551795			
;	GENERAL INFORMATION:			
;	APPLICANT: Marc J. Rubenfield et al.			
;	TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS			
;	TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS			
;	FILE REFERENCE: 107196_136			
;	CURRENT APPLICATION NUMBER: US/09/252,991A			
;	CURRENT FILING DATE: 1999-02-18			
;	PRIOR APPLICATION NUMBER: US 60/074,788			
;	PRIOR FILING DATE: 1998-02-18			
;	PRIOR APPLICATION NUMBER: US 60/094,190			
;	PRIOR FILING DATE: 1998-07-27			
;	NUMBER OF SEQ ID NOS: 33142			
;	SEQ ID NO: 23288			
;	LENGTH: 701			
;	TYPE: PRT			
;	ORGANISM: Pseudomonas aeruginosa			
;	US-09-252-991A-23288			
Query Match	15.8%	Score 379; DB 4; Length 701;		
Best Local Similarity	29.0%	Pred. No. 1.6e-21; Mismatches 67; Indels 22; Gaps 4;		
Matches	96;	Conservative		
Qy	158	EEAAAADDELVARLFPLMKLTFDQIAMDITYDSYAQRILHEDIDSROQIANAYATHYE	217	RESULT 11 US-09-252-991A-31412 ; Sequence 31412, Application US/09252991A ; Patent No. 6551795 ; GENERAL INFORMATION: ; APPLICANT: Marc J. Rubenfield et al. ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS ; FILE REFERENCE: 107196_136 ; CURRENT APPLICATION NUMBER: US/09/252,991A ; CURRENT FILING DATE: 1999-02-18 ; PRIOR APPLICATION NUMBER: US 60/074,788 ; PRIOR FILING DATE: 1998-02-18 ; PRIOR APPLICATION NUMBER: US 60/094,190 ; SEQ ID NO: 31412 ; LENGTH: 614 ; TYPE: PRT ; ORGANISM: Pseudomonas aeruginosa US-09-252-991A-31412
Db	384	QDIAGEGDILTKR---LATVSRDFEGVLDQAFNQFVEIHRSL---REVAGTA---	430	
Qy	218	PLSSLEATSDQVAERTDTMARTDDQVDRMADSREISSYSAVEVASTADDVRTSED	277	
Db	431	-HKLHDVQVNAQNSMANSDQSNNTNSVAAINEGAAQEIYANADASHHASD	488	
Qy	278	AEALAQGEAAADDALATMDIDEATDGVAGBOLGEAAADVSVTGVTGDIABQTNML	337	
Db	489	ANHQAEQDGKQVQEQTIRAMELSEKISASCANIAINSRTVNQIGLEVTKGISEQTNLL	548	
Qy	338	ALNASTEAAARAGEAGEGBFAVVADEVKALAESESRESTOREYELVQMOAETEETVDOLDEV	397	
Db	549	ALNAAIEAAAGEAAGRGAQVADEVNLAAQRAQESAAQIQMIEBLQVAREAVATMVE	608	
Qy	398	NORIGEVVERVEAMETLQLETITDAVEDDAASGMQEVTATEQAVSTEEVAAEMYDGVDDRA	457	
Db	609	QRYSLSEVTEANRABESLSVSTRERRIDGDMNQSYSTATEQOTA-----WVDSLNNDI	661	
Qy	458	GEIAALDDIADATDQVTRVEVRETYVGK	488	
Db	662	TEINTLNQEGVENLQATLRACGELETQAGSL	692	
RESULT 10				
;	Sequence 5172, Application US/09328352			
;	Patent No. 6562958			
;	GENERAL INFORMATION:			
;	APPLICANT: Gary L. Breton et al.			
;	TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER			
;	TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS			
;	FILE REFERENCE: GTC99-03PA			
;	CURRENT APPLICATION NUMBER: US/09/328,352			
Query Match	15.6%	Score 373; DB 4; Length 614;		
Best Local Similarity	26.6%	Pred. No. 3.8e-21; Mismatches 214; Indels 90; Gaps 18;		
Matches	142;	Conservative		
Qy	9	VTAQYRNGDGHALADRIG-----LDEAFAWLSPFG-IDDDTMALAEQPLPEA	59	
Db	108	LTDGVVR-----AYRELLGGPVRAAQIIDEA---NLQFRGQVQEWKNVLGRQ-----	152	
Qy	60	TADALVTDYDHLYSEY-TQDLFANSKTKEQLEKTAEYLGLGRHE-----DTEYAAQ	115	
Db	153	-AEA QTQXWSQFFAQERAQVQDILGRGSVAEGELKDRVERL---REHRRGTAYTQG	206	
Qy	116	RARIGKIHDPVGLGPDVYLG-----ATTRYTTGHLDAADDY-----VADR-	156	

207 RQRF-----LEAGADPIAGDQAVTGDIDRATTAAQMQLRDELHQASDLHSSSISAFARRT 260
 157 -----GEEAARAVDE-LVVARSL-PMLKLTFDQQIAMDITYIDSYAQRLLRBEIDS 203
 261 MLLGSVLIGASLAVALLSMVNRLNVRPQLHIAQLSHDFGERIEIRRDELGK 320
 204 RQEELANAVATHVEAPLSSLEATSQDVAYERDTD-MEARTDDQVDRMADSREISS 256
 321 LALANTRLDPLVIDFDRLRSTRDLSASSLNALASLMAGTREDFQRSDQVATAME 380
 257 VSASVREVASSTADDVERTSEADELAACQGEAAADDALATMDIDEATDGVTAGVEQLGER 316
 381 MSATAQEVARYAGDAARADEADDSAGRCMRKEIDHTVVEIQLESD 440
 317 AADVESUTGVDIILADQTMNLAINSLAEEAARGAESEGAUVADEKLAESREOSTRV 376
 441 SGRIGKVLDVTRGIAETQNLLALNAALEAARGDAGRGFAVVADETRLAQRTAESIAEI 500
 377 EELVEQMQAEETEETDQDLEVNQRIQGVVERVEEAMETLQETIDAVEAASGMQEVSSTAT 436
 501 HQLIDTVYANGAWAAAEGSOSRSLRAFGAEEVANQAMRLPITASVSDRMNRQATAA 560
 437 DEQAVSTVEAEMVGDYDVGTDRACEIAAALDDIADATDQVTRVEEVRETGVKLS 489
 561 EEQTAVADISRNLT-----EIAS---IASNOBQVEOTEAAASRDHLGLS 602

RESULT 13
 US-09-252-991A-31724
 ; Sequence 31724; Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenstein et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; SEQ ID NO: 31724
 ; NUMBER OF SEQ ID NOS: 33142
 ; LENGTH: 760
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-31724

Query Match 15 3%; Score 366; DB 4; Length 760;
 Best Local Similarity 31 %; Pred. No. 1-8e-20;
 Matches 97; Conservative 61; Mismatches 126; Indels 22; Gaps 4;
 Qy 195 QRL-HDBDSRQELA--NAVATHYAPLSSLEATSQDVARTDQVDRMADV 250
 Db 454 QRLPHTGRDGEELAGWNFRFLDKQIIRDVKVSYRDASTADSAISQTSACMQQ 513
 Qy 251 SREIISVSVASVEEVASTADDVTRTSEDAEAALAQCGEAAADDALATMDIDEATDGVTAG-
 Db 514 FREIDQVATASHEMATAQDVARSAAQAADARGAQATRGLALIDRTQISDIAANL 573
 Qy 309
 ; Query Match 15 4%; Score 368.5; DB 4; Length 573;
 Best Local Similarity 25.0%; Pred. No. 7.9e-11; Gaps 15;
 Matches 129; Conservative 89; Mismatches 175; Indels 123; Gaps 15;
 Qy 310 -----VEOLGERADEVSTGVIIDIAEQTNMLNNASEAARGAGEGFAVVADEVK 363
 Db 574 TSAMGQEVQLAASSSERFIGSLEVIRAIAEOTNLLALNAATEAACRGDFAVADEVR 633
 Qy 364 ALAEESREQSTRVTEELVEMQAEETEETVQDLEVNQRIQGVVEREEAMETLOITDAVE 423
 Db 634 NLARRTQDSYEQIRGVIEGLQQGTDDVDMHGSTHRQAOQSVEQDVBAVALQRIGEAVT 693
 Qy 424 DAASGNGQEVSTATDQAVSTVEAMVGDYDVRGSEIAALDDIADATDQVTRVVEVRE 483
 Db 694 VINDMNLIQIAAEEOSVVAEINRN-----AIRDVTESLSSOAEEAQVSQ 742
 Qy 484 TVGKLS 489
 Db 743 SLNRLLA 748

RESULT 14
 US-09-252-991A-25899
 ; Sequence 25899; Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:

Db 216 EAPLSSLEATSDQVAERTDQVDRMADV-----VSREISVSVASVEVAST 267
 Qy 184 IAMDTY-----IDSYAGRLHD-----EDTSRQE----LANAVATHV 215
 Db 292 QPLIREVGRTVGEVADSAGSLAGMTAAN-DRLINSBHAQDQVSTAATQMSAVHEVARN 350

APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196-136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1999-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-02-18
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO: 25899
 LENGTH: 613
 TYPE: PRT
 ORGANISM: *Pseudomonas aeruginosa*
 US-09-252-991A-25899

Query Match 15.0%; Score 359; DB 4; Length 579;
 Best Local Similarity 25.0%; Pred. No. 4.4e-20;
 Mismatches 85; Indels 114; Gaps 15;
 Matches 133; Conservative 85;

Qy 44 DDTMAAALAAEQLPFEATADALYTFDYLLEY--ERTQDLFANSTKTEVQLKETQAYL 100
 Db 70 DNTYAA-----NTLRTDIAQAQFSYTAQPOVED--ANITFIQTMDARKA-- 113
 Qy 101 LGIGRGEYDT--BYAAQARARTGKIHVLGLGPDVILGAYTRYTGLI-- 145
 Db 114 -----EDTLMNBYEVLSAQTG--FDQORMAFIKYOKSRYDLDLVQYLNERNNTNAI 165
 Qy 146 -----DALADDVYADRGEAAADEVDEL-----VARFLPMKL 178
 Db 166 RLILKSPANDNFTNSLYATDFITDDVNPSVTEADNYHEMILWIASLFGMGLIFTTLV 225
 Qy 179 TF-----DQQIANDTYIDS-YAQRHLDEIDSQEL----ANAVATHVEAPLS 220
 Db 226 LFWMRKHIIVRINQMIQEYQATASGDLISRIDHIDTRNEIDOLMLGILQMRARLKMVS 285
 Qy 221 SLEATS-----QDVAERBTDMRARTDDQVDEMADSREISSVVASYEVASTDYVRR 273
 Db 286 AIRNSSTTIVSGYQEIAAGNNDISRTEQASALEETASSMEGLATVRNNTESAREVTQ 345
 Qy 274 TSDEAELAQQGEAAADDALATMTDIDEATDGTVAGEQQLGERAADYESVTGVIDDIAEQ 333
 Db 346 LMNSTADIAVQGEGHSNKNMVMIDI-----ADESQKCEITAVIDDAFQ 391
 Qy 334 TNMLALNASIEAARAGAGEGFIAVVADEVKALAEESREQSTRVBEYLVQMOAETTEETVDQ 393
 Db 392 TNILALNAVEAARAGEOGRGFSVVASEVRNLAQRSAZAAKEIKELIBSTILRVKGNDL 451
 Qy 394 LDEVNQRIGEGVERVEAMETOITDAVEDAASGMQEVSATDEQAVSTTEEVAEMYDGV 453
 Db 452 VEQVSLSMGEIVTSVNVHVGSMREBILASEQTRGIAQTSLAVNMEDKATQONAAMV--- 508
 Qy 454 DDRAGEIAARALDDIADATDQQVRT--VEVRRET 484
 Db 509 -EQSTAASATLSEQASLDMIVNTFKVEQVAVT 540

Search completed: August 10, 2004, 15:30:19
 Job time : 26.8644 secs

RESULT 15

US-09-543-681A-6665
 Sequence 6665, Application US/0943681A
 Patent No. 6605709
 GENERAL INFORMATION:
 APPLICANT: GARY BRETON
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 FILE REFERENCE: 2709_1002-001
 CURRENT APPLICATION NUMBER: US/09/543,681A
 CURRENT FILING DATE: 2000-04-05
 PRIOR APPLICATION NUMBER: US 60/128,706
 PRIOR FILING DATE: 1999-04-09
 NUMBER OF SEQ ID NOS: 8344
 SEQ ID NO: 6665
 LENGTH: 579
 TYPE: PRT
 ORGANISM: *Proteus mirabilis*
 US-09-543-681A-6665

Hunt
Bart

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 10, 2004, 15:28:46 ; Search time 71.0412 Seconds
(without alignments)
2159.184 Million cell updates/sec

Perfect score: US-09-455-978B-2
Scoring table: BL050N62 Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match %

Listing first 45 summaries

Published Applications AA:*

```
1: /cgn2_6_ptodata/2/pubpaas/US07_PUBCOMB.pep:*
 2: /cgn2_6_ptodata/2/pubpaas/PCT_NEW_PUB.pep:*
 3: /cgn2_6_ptodata/2/pubpaas/US06_NEWT_PUB.pep:*
 4: /cgn2_6_ptodata/2/pubpaas/US05_PUBCOMB.pep:*
 5: /cgn2_6_ptodata/2/pubpaas/US07_NEW_PUB.pep:*
 6: /cgn2_6_ptodata/2/pubpaas/PCTCUS_PUBCOMB.pep:*
 7: /cgn2_6_ptodata/2/pubpaas/US08_NEW_PUB.pep:*
 8: /cgn2_6_ptodata/2/pubpaas/US09A_PUBCOMB.pep:*
 9: /cgn2_6_ptodata/2/pubpaas/US09B_PUBCOMB.pep:*
10: /cgn2_6_ptodata/2/pubpaas/US09C_PUBCOMB.pep:*
11: /cgn2_6_ptodata/2/pubpaas/US09_NEW_PUB.pep:*
12: /cgn2_6_ptodata/2/pubpaas/US10A_PUBCOMB.pep:*
13: /cgn2_6_ptodata/2/pubpaas/US10B_PUBCOMB.pep:*
14: /cgn2_6_ptodata/2/pubpaas/US10C_PUBCOMB.pep:*
15: /cgn2_6_ptodata/2/pubpaas/US10D_PUBCOMB.pep:*
16: /cgn2_6_ptodata/2/pubpaas/US10_NEW_PUB.pep:*
17: /cgn2_6_ptodata/2/pubpaas/US60_NEW_PUB.pep:*
18: /cgn2_6_ptodata/2/pubpaas/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	399.5	16.7	535	12	US-10-282-122A-66353
2	399.5	16.7	535	12	US-10-389-647-531
3	386.5	16.1	682	12	US-10-282-122A-66174
4	374.5	15.6	680	12	US-10-282-122A-69670
5	371.5	15.5	686	12	US-10-282-122A-68152
6	361.1	15.1	644	12	US-10-282-122A-77591
7	359.5	15.0	891	9	US-09-272-809-5
8	356.5	14.9	539	12	US-10-282-122A-51025
9	348.5	14.6	679	12	US-10-389-647-372
10	348	14.5	626	12	US-10-282-122A-77066
11	338	14.1	686	12	US-10-282-122A-52305
12	332	13.9	547	12	US-10-332-288-78
13	326	13.6	641	12	US-10-282-122A-77278
14	315	13.2	643	12	US-10-282-122A-77602
15	314.5	13.1	501	12	US-10-282-122A-69564

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Sequence 77047, A
Sequence 76517, A
Sequence 34, APP

Sequence 34, APP

Sequence 34, APP

Sequence 6156, APP

Sequence 6157, APP

Sequence 462, APP

Sequence 460, APP

Sequence 454, APP

Sequence 450, APP

Sequence 449, APP

Sequence 448, APP

Sequence 447, APP

Sequence 446, APP

Sequence 445, APP

Sequence 444, APP

Sequence 443, APP

Sequence 442, APP

Sequence 441, APP

ALIGNMENTS

RESULT 1
US-10-282-122A-66393

; Sequence 66393, Application US/10282122A

; GENERAL INFORMATION:

; Publication No. US20040029129A1

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheyly

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

FILE REFERENCE: ELTRA 034A
CURRENT APPLICATION NUMBER: US/10/282122A

PRIOR APPLICATION NUMBER: 60/091,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06

PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/267,636

FILE REFERENCE: ULTRA 034A
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 Remaining Prior Application data removed - See File Wrapper or PALM.

SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 66393
 LENGTH: 535

TYPE: PRT
 ORGANISM: *Pseudomonas aeruginosa*
 US-10-389-647-531

Query Match 16.7%; Score 399.5; DB 12; Length 535;
 Best Local Similarity 25.2%; Pred. No. 1.8e-18;
 Matches 132; Conservative 111; Mismatches 206; Indels 75; Gaps 16;

Query Match 16.7%; Score 399.5; DB 12; Length 535;
 Best Local Similarity 25.2%; Pred. No. 1.8e-18;
 Matches 132; Conservative 111; Mismatches 206; Indels 75; Gaps 16;

Qy 3 NDNDTLVTDYRNGIDGHALADRIGLEDAETAWRLSFTGIDDNTMALLAEEQPLFEATAD 62
 Db 37 SENELSVNA-LRNHMEGDMHHD-----ALRADVLAFAFVY-QPGDGAEE 78

Qy 63 ALVTDFYDHLESVERTQDLFANSTKTYEQ----LKEHQAYLLSLIGRGRGYDTEAAQRA 117
 Db 37 QVRQDQLQEHSMWFR-----KVEQNQGLPINDATHQALVEL-RPDELEYIGAEE 127

Qy 118 RIGK-IHDVLGLGPDV-YLGAYTRYTGTGLDADY-159
 Db 79 QVRQDQLQEHSMWFR-----KVEQNQGLPINDATHQALVEL-RPDELEYIGAEE 127

Qy 118 RIGK-IHDVLGLGPDV-YLGAYTRYTGTGLDADY-159
 Db 128 IVGKAUDDPAARAELPQFVQAF-KELBEGNEQNSLIEKHBEQTNRAREDSMRYSAWML 186

Qy 160 -----AAAVADELVARFL----PMKLKLTEDQOJAMDTY----IDS-YAQRLHDEI 201
 Db 187 AGGILVACLVQGLCQLRQLRQVLPDKLRKVSSARTIAQGNLQEPVGVDNSDEAGQLQRAL 246

Qy 202 DSROELANAVATHVAPLSSLEATSQDAERTDTMRARTDDQVDEMADVSREISSVSASV 261
 Db 247 GEMQENLQRQMTTIIQESSEELHDSOSIGTSQSIVHGSQQADSATSMASMPMINTNI 306

Qy 262 EEVASTADDVRRTSRDAEALAQOGGAAADDALATMDIDEDATGTVAGEQLGERAADVE 321
 Db 307 SQISDHADNARVISAKSEELASSGGQVINVVEGRNSRADIUVNQSTS-TALGQSDEIH 366

Qy 322 SVTGVIDDIAEQTNNMLNASTIAARAGEAGEGFNADEVKALAEESPEQSTVVEELVE 381
 Db 367 SRIQITKGTAEQTNLALNATEARAGEAGRGFVAVADEVRGLHLARTTOSTQEITAMIE 426

Qy 382 QMOAETEETVDQLDENVNRQIGEVVERVEAMETLQEITDAVEDASGMQEVSSTATEQAV 441
 Db 427 RIRASTGQANSMSMAGVSRYNEGGSFARAGVSNEBILUQTRHAAVSVDIESQTRREQSR 486

Qy 442 STEVAEMVGDVDDRAGEIAALDDIADATDQQVTRVEEYRET 485
 Db 307 SQISDHADNARVISAKSEELASSGGQVINVVEGRNSRADIUVNQSTS-TALGQSDEIH 366

Qy 322 SVTGVIDDIAEQTNNMLNASTIAARAGEAGEGFNADEVKALAEESPEQSTVVEELVE 381
 Db 367 SIIQVTKKGIAEQTNNLALNAAIEARAGEAGRGFVAVADEVRGLAARTOSTQEITAMIE 426

Qy 382 QMOAETEETVDQLDENVNRQIGEVVERVEAMETLQEITDAVEDASGMQEVSSTATEQAV 441
 Db 427 RIRASTGQANSMSMAGVSRYNEGGSFARAGVSNEBILUQTRHAAVSVDIESQTRREQSR 486

RESULT 2 STEEVAEMVGDVDDRAGEIAALDDIADATDQQVTRVEEYRET 485

Qy 442 STEEVAEMVGDVDDRAGEIAALDDIADATDQQVTRVEEYRET 485
 Db 487 ASDETAQRVELIAQSQQNTQAHMNA-AT---ARLNVEAATM 526

RESULT 2 STEEVAEMVGDVDDRAGEIAALDDIADATDQQVTRVEEYRET 485
 US-10-389-647-531
 Sequence 531, Application US/10282122A
 Publication No. US200400291129A1

GENERAL INFORMATION:
 APPLICANT: GREENBERG, E., Peter
 APPLICANT: SCHUSTER, Martin
 APPLICANT: LOSTROH, Candi
 TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
 FILE REFERENCE: U12-038CP
 CURRENT FILING DATE: 2003-03-14
 PRIOR APPLICATION NUMBER: 09-653730
 PRIOR FILING DATE: 2000-09-01
 PRIOR APPLICATION NUMBER: 60/153022
 PRIOR FILING DATE: 1999-09-03
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 531
 LENGTH: 535

RESULT 3 Sequence 66174, Application US/10282122A
 Publication No. US200400291129A1
 GENERAL INFORMATION:
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forzych, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELTPA-034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335

PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SEQ ID NO: 66174
 LENGTH: 682
 TYPE: PRT
 ORGANISM: *Pseudomonas aeruginosa*
 US-10-282-122A-66174

Query Match 16.1%; Score 386.5; DB 12; Length 682;
 Best Local Similarity 27.5%; Pred. No. 1.8e-17; Mismatches 77; Indels 33; Gaps 4;

Matches 95; Conservative 77; Mismatches 140; Indels 33; Gaps 4;

US-10-282-122A-66174

Query Match 15.6%; Score 374.5; DB 12; Length 680;
 Best Local Similarity 24.2%; Pred. No. 1.1e-16; Mismatches 103; Indels 71; Gaps 12;

Matches 120; Conservative 103; Mismatches 201; Indels 71; Gaps 12;

US-10-282-122A-66170

Query Match 15.6%; Score 374.5; DB 12; Length 680;
 Best Local Similarity 24.2%; Pred. No. 1.1e-16; Mismatches 201; Indels 71; Gaps 12;

Matches 120; Conservative 103; Mismatches 201; Indels 71; Gaps 12;

US-10-282-122A-66170

TYPE: PRT
 ORGANISM: *Pseudomonas syringae*

US-10-282-122A-66170

Query 23 ADRIGLDEAFIA----ATADALVDF 68
 Db 207 ADAFGRDASQFGRVLNGMLEGNATLRITQVEDRARLAELAEFVSGSVDEILETS 266

Qy 69 YDHLESYERTQDLPANSTKTVBQLKETQAEYLGLRGEXYDTEAAQRARIKHD-VLG 127
 Db 267 PEYQREFRSQNT-NTSQTLDETSVANSI----ENIARKPTMTYGGVILG 315
 Qy 128 LGPDVYLGAYTRYYTGILDALADYVDRGEAAAADVDELVARFLPLKLITPFDQIA-- 185
 Db 316 L----LALMSIIILGLV----MVRINRNQLRBTAKSERNOTAIMLLDETENLADG 364
 Qy 186 -----MDTYIDSQAQLHDEIDSQRLA---NAVATHVEAPLSSLEATSQDAERTD 234
 Db 365 DLTTVATSTEDFTSALADSINYSDQRLRVVTNTLAEQVASVETOTAMOL----- 420

Qy 235 TMRARTDDQVDRMDSRBISSYSASVEEVASTDVRRTSEDAEALAQGEAAADDALA 294
 Db 421 ---AASEHQALQSAASTAVNDMABASIDQSYANASESSASAERSVATANKNEYVQNTTH 477

Qy 295 TMTDIDEADTGVTAGVEQLGERAADVESVTGVIDDIAEQTNMLALNASTEAARAGEAGG 354
 Db 478 GMDNIREQIQTDTSKRKIRKGESSOEIGDIVSLDDIAQTNILNALAQSMAGDGRG 537

Qy 355 FAVVADEVKAEEFSRQEOSTRVEELVEMOQAETETVQDOLDEYNQIGFQVEREAMENT 414
 Db 538 FAVVADEVRAEFSRQEOSTRVEELVEMOQAETETVQDOLDEYNQIGFQVEREAMENT 414

Qy 415 LOEITDAVEDAAGSMQEESTATEQAVSTEEVAEMDGVDRAJEIAALDDIADATQQ 474
 Db 598 LGEGEGSRVLAELIESTDAAHQAES-----AQGQSTMVVIQTSQZ 643

Qy 475 VRTVEEVRETVGKLS 489
 Db 644 TSGTSATAESIGNIA 658

RESULT 4
 US-10-282-122A-66170
 Sequence 63670, Application US/10282122A
 Publication No. US20040029129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Hasebech, Robert
 APPLICANT: Ohlsen, Mari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Travick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITTA-03A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078

RESULT 5
 US-10-282-122A-68162
 Sequence 68162, Application US/10282122A
 Publication No. US20040029129A1

Query Match 14.9%; Score 356.5; DB 12; Length 539;
 Best Local Similarity 26.6%; Pred. No. 1.3e-15; Gaps 15;
 Matches 126; Conservative 81; Mismatches 153; Indels 113;

Qy 38 SFTGIDDDTMAA-----AEQPFLFEATADALTYDFYDHLSEYRTQDLFANSTKTVEQL 92
 Db 126 SPAASLDNEAALERNDAYLQLKVNKGQASAAFAARLGEFSTNLKISSETLAAHT 185
 Qy 93 KETOAEYLGLGRGEYDTEYAAQPARIGKIHVDVGLGPDVYLGAYTRYTGGIADDALANDV 152
 Db 186 RETIMLYV-----YAA-----LILLIVCIAGSYLYMTRAVRPLQR-- 222
 Qy 153 VDRGEAAADEVLAFLVPLMLKLTDQDQIAMDITYDSYAQLHDEIDSROBLANVA 212
 Db 223 -ASQQFERNAGD-LSARI-----EGNSRNEIGAILT 252
 Qy 213 T--HVEAPLS-----SLEATSDQVAERTDTMARTDDQVDRMADVSPREISSVSA 259
 Db 253 ALRHMQESLRTVAAVRGVDEINVGSREISAGNTDLSSRTEQ----AASLEETAA 305
 Qy 260 SVEEVAST---ADDVRETSF---DAELAACQCBAAAADDALATMDIDEATDGVTAGEQ 312
 Db 306 SNEQLASTVKONADQANOLAASASVVAESGGSAVSVSTM-----DGISASSRK 358
 Qy 313 LGERRAADVESVTGVID1IAEOTNMLAANASIEARAGFAGEGFPAVVADEVKAIAEESREQ 372
 Db 359 ISE-----IVSVIDGIATQFQTNILANAAVEARAGQKGPAVVAEVRSIAQRAA 411
 Qy 373 STRVEELVEMQMQAETEETVQDQLEVNQRIGEYVERVEAMETLQEITDAEADASGMQE 432
 Db 412 ARCKIVLVE-----DSVD----KVGTGSSQVERAGATMQEITVAVSRVRTDIMG 457
 Qy 433 STATDEQAVSTEEVAMTDGVDDRAGEITAALUDDIADATDQVVRTBEVRETV 485
 Db 458 SAASEE0SGTEQVNTRAVSQMDDEVQTMQAALVEEAAAGSLQEQAORLAAEV 510

RESULT 9
 US-10-389-647-372
 ; Sequence 72, Application US/10389647
 ; Publication No. US20040033549A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GREENBERG, E. Peter
 ; APPLICANT: SCHUSTER, Martin
 ; APPLICANT: LOSTROH, Candi
 ; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
 ; FILE REFERENCE: UTZ-038CP
 ; CURRENT APPLICATION NUMBER: US/10389647
 ; PRIOR APPLICATION NUMBER: 2003-03-14
 ; PRIOR FILING DATE: 2003-09-14
 ; PRIOR APPLICATION NUMBER: 60/153022
 ; PRIOR FILING DATE: 1999-09-03
 ; NUMBER OF SEQ ID NOS: 710
 ; SOFTWARE: PastSeq for Windows Version 4.0
 ; SEQ ID NO: 372
 ; LENGTH: 679
 ; TYPE: PPT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-10-389-647-372

Query Match 14.6%; Score 348.5; DB 12; Length 679;
 Best Local Similarity 25.9%; Pred. No. 5.9e-15; Gaps 14;
 Matches 123; Conservative 81; Mismatches 178; Indels 95;

Qy 20 HALADRIGDEAEI---AWRLSFGTGDIDDTMAAALAEEQPLFEATADAL--VTDFYDHLES 74
 Db 242 HLLANLTGPHAEILNGGRFLPSLVDV-----EVENDANERLGSQVWTDRLTE 289
 Qy 75 YERTQDLFANSTKTVEQLKETQAELLGJRGEDTEVAQRARIGKIFHDVLGJGPDVY 134
 Db 290 HRAEQE-----VSQVQAAA-----GDPSKRVEEAGKEGGFFKLAKD-- 327

Qy 135 GAYTRYTGTGIDDLADDVVADRGEEAAADEVLFPMKLULTDQQIAMDTYIDSVA 194
 Db 328 -----LNSLVD--TADRGIRD-----VSRMLGALAQCDLTQPLEAD-YQTFPG 367
 Qy 195 QRLAIDEDBSRQELNAVATHVEAPLSSLEATSQVAERDFTMRARTDDQVDRMADVSRBT 254
 Db 368 QLKFSNDTAQSLSRMLGQIREA-ADTINTAASEIASGNAELSRATEQQASSLETTASM 426
 Qy 255 SSVASVVEVASTADVERTSADDEALAOQGEAAADDALATMDIDEATDGVTAGVEQIG 314
 Db 427 EELTSTVKLNAAENRQANSLAANASSEVATGGTVVQRKVSTMSINES-----474
 Qy 315 ERADVESVTGVID1IAEOTNMLAANASIEARAGFAGEGFPAVVADEVKAIAEESREQ 374
 Db 475 --ARKIAD1IGVIGDIAFQTNILANAAVEARAGFAGEQCRGFAVAGEYVTLAQRSAAAK 532
 Qy 375 RVEELVEMQMQAETEETVQDQLEVNQRIGEYVERVEAMETLQEITDAEADASGMQE 434
 Db 533 EIKLILS-----DSVDKVE----NGNTLVAAQGOTMSDIVVAVRRTDIMSSETAA 578
 Qy 435 ATDQXANSTEEVAMTDGVDDRAGEIAALDDIADATDQVVRTBEVRETVKLIS 489
 Db 579 ASAEQSTGIEEVNSAVSQMDMTQQNALVEEAAA-----AEAMQEQAQLIN 626

RESULT 10
 US-10-282-122A-77066
 ; Sequence 77066, Application US/10282122A
 ; Publication No. US2004002912A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cherry
 ; APPLICANT: Haslebeck, Robert
 ; APPLICANT: Ohlsken, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forayth, R.
 ; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA-034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-05-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78514
 ; SOFTWARE: Patent In version 3.1
 ; SEQ ID NO: 77066
 ; LENGTH: 626
 ; TYPE: PRT
 ; ORGANISM: Vibrio cholerae

US-10-282-122A-77066
Query Match 14.5%; Score 348; DB 12; Length 626;
Best Local Similarity 27.4%; Pred. No. 5; 7e-15;
Matches 92; Conservative 78; Mismatches 134; Indels 32; Gaps 6;
SEQ ID NO: 52305 ; LENGTH: 686 ;
TYPE: PRT ;
ORGANISM: Clostridium botulinum
US-10-282-122A-52305

QY 174 MLKLTFDQQI-----AMDTYIDSAYAQLRHLDEIDSQELANAVATHEAPLSS 221
Db 298 MSPKTLDEAKIDASGGGDLTRKDIDKEFSEALQNSFTMELSQIRQKTIASG 357
Qy 222 LEATSDQVAERTDMARTDDQVDRMDSREISSVSASVEEYASTADDVRRTESEDAL 281
Db 358 VLDGAEKTAANEAEYSLRVYEQQL-----QELEQULATANEMANTASVANSAQVAADA 410
Qy 282 AQQEBAADDAALATMDIDATDGTW-----AGVEOLGE--RAAD-VESTVYIDDA 331
Db 411 AKEGESA---SLEGSVVUHETTDALQRSLTRIGSSVEDYKELVKATDRIFTEVDINDIA 467
Qy 332 EQTNMLAINASITEARAGEAGEGFAYVADFKVALAESRFOSTROSTVEELVQMQATEETIV 391
Db 468 DOTNLLNALNATEARAGEAGEGRGFAVADTQVQSTMEQVSTADEQVSTEEVAEMYD 451
Qy 392 QDLDEVNRQGEVVERVEEAMETOIITEDAAGSMQBVSTATDEQAVSTEEVAEMYD 451
Db 528 RSMDESKLTDIVVETKTQNQEYNEKISLVQPAUTHRISDMNLQIASAAEQSLVABEINNNTV 587
Qy 452 GVDDDEGEIAAADDIADATDQQVTVETRETVGK 487
Db 588 NIKDLISKLSEAASNAGTENNAQSVKVKQEQNELLNE 623

RESULT 11
US-10-282-122A-52305
Sequence 52305, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cherry I
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA_034A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 52305
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Clostridium botulinum
; Query Match 14.1%; Score 339; DB 12; Length 686;
; Best Local Similarity 26.8%; Pred. No. 3e-14;
; Matches 91; Conservative 80; Mismatches 130; Indels 38; Gaps 4;
; Query Match 178 LTFFOQIAMDTYIDSAYAQLRHLDEIDSQELANAV-----ATHVEAPLSSLE----ATSQ 227
; Db 350 LTERKVYVNEDEIGKLSKLFNTMDSREITRINNFSQLASSAEOTSVAVE 409
; Qy 228 DVAERTDTMRARTDDQVDRMDSREISSVSASVEEVASTADDVRRTEDAALAQGEA 287
; Db 410 EISSATEEEIASGAENQVKASNNESSLLNDUMGMNYTKEFDEIISFSNTNTLASKGQE 469
; Qy 288 ADDALATMDIDEATDGVTAGYEQQLGERAADVESVTGYIDDIAEQTNMLANASIEAR 347
; Db 470 NMGNMVQONATIONSVNNSNNIMYDLQRNSEEONIVETINTADQTNLLANASIEAR 529
; Qy 348 AGEAGEGFAYVADYKVALAESRFOSTROSTVEELVQMQATEETIVDQDLEVNQIGEYVER 407
; Db 530 AGEAGEGFAYVADYKVALAESRFOSTROSTVEELVQMQATEETIVDQDLEVNQIGEYVER 589
; Qy 408 VEEAMELTQEI-----TDAVEDAASGMQEVTSTATDEQAVSTEEV 446
; Db 590 VAEVKESLGELINSFSNVHKFAVSUDSMITASNDTSITANASKLYDIEITISNTASANTEEV 649
; Qy 447 AEMVNGVDDRAGETAALDDIADATDQYRTVEEYRTV 485
; Db 650 AASTB-----EQSATIEELTESIEKLVSMVENLKESV 681

RESULT 12
US-10-332-288-28
Sequence 28, Application US/10332288
Publication No. US20040054165A1
GENERAL INFORMATION:
APPLICANT: RAINEX, Paul Barton
APPLICANT: SPIERS, Andrew Julian
APPLICANT: BANTINAKI, Elena
TITLE OF INVENTION: BACTERIAL POLYSACCHARIDE AND BIOFILM DEVELOPMENT
CURRENT APPLICATION NUMBER: US/10/332,288
PRIORITY APPLICATION NUMBER: PCT/GB01/03077
PRIORITY FILING DATE: 2003-10-06
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 28
LENGTH: 547
TYPE: PRT
ORGANISM: Pseudomonas fluorescens
US-10-332-288-28

Query Match 13.9%; Score 332; DB 12; Length 547;
Best Local Similarity 25.0%; Pred. No. 5.6e-14;
Matches 119; Conservative 84; Mismatches 189; Indels 84; Gaps 12;

Qy 66 TDFYDHLEYERTQDLFANSTKTEVQLRETQAEYLGLGRGEYDTEVAA-----114
Db 89 TDKNDKSFEARLEIQMANYKTI----HGQAD----RMEFDNFKAHINYNKVLAQV 138

Qy 115 -QARAGKIHVDLGLGPDVYLGATRYTGYJDLAADDV-YADRGEEAAAAYDELV---168
Db 139 LERVEANDLPGANQNLBEQLPTIWTEGRMKUNDITENKVNSDR--ATAADEAVLSAK 195

QY 169 -----ARFLPMLKLLFDQQIAMDVTID-SYAQRHLH---DB 200
 Db 196 ISMAVSLILAILAAGLGGILLMRAIMAPRIVDI----LETRNDGDSKRNLERD 250
 QY 201 IDSRDELLANAVATHYEPLSSLEATSQDVARTDDQVDRMADVSEISSVASS 260
 Db 251 FGAVETGGNDMMTEITALVISQAQRSSVQVTTSVTEIAITSQQQATAETTAETGAT 310
 QY 261 VEEVASTADDVVRTSEDAFALAOGEAAA--DDALA---TMDIDFATGVTTAGVSQI 313
 Db 311 SREAAATSQDLVRMTTEVSTAADASVAAGSGQGLRNBTETHVSMGADLYNAKLIL 370
 QY 314 GERAADVSEVTGVIDDIAROTNMALNAAStEARAAGEAGEFAVVADEVKALAEESRPOS 373
 Db 371 NEKAGNINGOYVVTIVKAADOTNLISLNAAStEARAAGEGRGPAAVATVRRADQTAAT 430
 QY 374 TRVBEELVQMOAQATEETVTDOLDEVNQRGEGVERVEEMETLQBITDAVEDDAASGMQEVs 433
 Db 431 YDIEQMVRTEIQSAYSAGTMGMDKFSEEVTRGMFPEVQQGEQLSQTIHQVQALAPRVLMVN 490
 QY 434 TAIDEQAVSTEEVAEMVDGVDRAEIIAALDDIIATDQOQRTVEFRETYGKLS 489
 Db 491 EGMQAQATGAEQ1NH----ALVOLGDASSQTVESLRQASFADELS 532
 QY 532 -----ATMDIDE--ATDGVTAGVEOLGE 315
 Db 593 GTTSLTQTEQGARLGKERAESSLKACERKMLTGIEENKTIVTDEMILATTESTXAAATENLINE 452
 RESULT 13
 US-10-282-122A-77278
 Publication No. US20040029129A1
 GENERAL INFORMATION
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cherry
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA_034A
 CURRENT APPLICATION NUMBER: US/10/282.1122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/1191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-16
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-13
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR FILING DATE: 2001-02-16
 REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 77278
 LENGTH: 641
 TYPE: PRT
 ORGANISM: *Vibrio cholerae*

US-10-282-122A-77278
 Query Match 13.6%; Score 326; DB 12; Length 641;
 Best Local Similarity 24.1%; Pred. No. 1.7e-13;
 Matches 117; Conservative 97; Mismatches 164; Indels 108; Gaps 16;
 QY 45 DTWALAAEFQPLPFATADALVTDFDHTLESYERTQDLFANSTKVEQLKETQAE---- 98
 Db 203 DTVTLAETKAERNANIESLIKQISNGAE----MYSNGVKSSQSAESEYQKGHVN 255
 QY 99 YLLCIGRGRHYDTEYAAQRARIGKHDVGLGPDVYLGYTRYTGLLDAL-ADDVVAORG 157
 Db 256 KQTLIRNE----MDLVKD----SIVTLKADEMSSAAVALKSESEVSQAA 300
 QY 158 EEAAPAAVDELVARFLPMKLLTFDQIAMIADTYDSYARLHDIDTSROELNAVATHVEA 217
 Db 301 QEQAACE----SUKSLD-QQQIALDGAVTA-AQSLDLTELR----- 339
 Db 218 PLSSLEATSQDVAERTDTMRARTDDQVDRMADSREISSVSASVEEVASTADDVRRTSE- 276
 Db 340 -----TSTDIVKSSERVAAAREELSAGIEINRSSIMEGALNQISSGAHMAKSVE 392
 Db 393 GTTSLTQTEQGARLGKERAESSLKACERKMLTGIEENKTIVTDEMILATTESTXAAATENLINE 452
 QY 316 RAADVESUTGVDDINE-----QTMMLAINASIEARAGIAEGPAVVADEVKALAAE 368
 Db 453 -MANIERUSROLDKVIGGISNVIOTMLAVNGAVEARAGEYKGFAVVSSTDQNLLAND 511
 QY 369 SREQSTRVVEELVEMQMQAE-----TEE----TVDOLDEVNRONGEGERV 407
 Db 512 AAFNAEFTIKDQVNQIQQINIVRKDLADLSTVMEEQKAALTIKQDUNVRMDSVVLG 571
 QY 408 VEEAMETILOQETIDAVEADAASGMOEVSTATDEQAVSTEEVAEMVDYDVRAGEBIAALDDI 467
 Db 572 SKRISSEAGGTERSADARAGNQIAAESSHATEGAAATAARQSSSTSPLASATENI 631
 QY 468 ADATDQ 473
 Db 632 AAVADE 637

RESULT 14
 US-10-282-122A-77602
 Sequence 77602, Application US/10282.122A
 / Publication No. US20040029129A1
 / GENERAL INFORMATION:
 / APPLICANT: Wang, Liangsu
 / APPLICANT: Zamudio, Carlos
 / APPLICANT: Malone, Cherry
 / APPLICANT: Haselbeck, Robert
 / APPLICANT: Ohlsen, Kari
 / APPLICANT: Zyskind, Judith
 / APPLICANT: Wall, Daniel
 / APPLICANT: Trawick, John
 / APPLICANT: Carr, Grant
 / APPLICANT: Yamamoto, Robert
 / APPLICANT: Forsyth, R.
 / APPLICANT: Xu, H.
 / TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 / FILE REFERENCE: ELITRA_034A
 / CURRENT APPLICATION NUMBER: US/10/282.122A
 / CURRENT FILING DATE: 2003-02-20
 / PRIOR APPLICATION NUMBER: 60/1191,078
 / PRIOR FILING DATE: 2000-03-21
 / PRIOR APPLICATION NUMBER: 60/206,848
 / PRIOR FILING DATE: 2000-05-23
 / PRIOR APPLICATION NUMBER: 60/207,727
 / PRIOR FILING DATE: 2000-05-16
 / PRIOR APPLICATION NUMBER: 60/230,335
 / PRIOR FILING DATE: 2000-09-06
 / PRIOR APPLICATION NUMBER: 60/230,347
 / PRIOR FILING DATE: 2000-09-09
 / PRIOR APPLICATION NUMBER: 60/242,578
 / PRIOR FILING DATE: 2000-10-13
 / PRIOR APPLICATION NUMBER: 60/253,625
 / PRIOR FILING DATE: 2000-11-27
 / PRIOR APPLICATION NUMBER: 60/257,931
 / PRIOR APPLICATION NUMBER: 60/267,636
 / PRIOR FILING DATE: 2001-02-09
 / PRIOR FILING DATE: 2001-02-16
 / REMAINING PRIOR APPLICATION data removed - See File Wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 78614
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO: 77278
 / LENGTH: 641
 / TYPE: PRT
 / ORGANISM: *Vibrio cholerae*

PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: Patentin version 3.1
 SEQ ID NO: 77602
 LENGTH: 643
 TYPE: PRT
 ORGANISM: *Vibrio cholerae*
 US-10-282-122A-77602

Score 315; DB 12; Length 643;
 Query Match 13.2%; Best Local Similarity 24.9%; Pred. No. 9.2e-13; Matches 100; Conservative 77; Mismatches 167; Indels 58; Gaps 10; Score 315; DB 12; Length 643;

Query 133 YLG-AYTRYTGGLDALARDD-----VVAADRGEAAAAAYDELVAR-----170
 Database 249 YLGIEFDGYKRVK-NLTDDMYWTLMFYDKA-TALADYDVAVANAITGIVLIGISGV 305

Query 171 -----FIPMLKLTFDQIAMDITYDSQAQLHEDIDSQELAN-----209
 Database 306 LIFLINQIYPLRUKVAVIDLQGS--GLTRRL--EVNGNDLIAQISGFNRFSGNLQ 361

Query 210 ----AVATHYAPLSSLTEASQDAERTIDMRARTDDQVRMADSRETSVSASVEYA 265
 Database 362 NMMLQISDQQIISSSIEQSQ-TAKENEMLISHSETQ--VVTATQMSSEARTV 417

Query 266 STADDVVRTSSEDAELAQOGEAAADDALATMTDIDEATDGTVTAGVBQLGRRAADESYSTG 325
 Database 418 ESVTQNQITPAASEKAQQLIVTNAVSVTSLINDVENSSESINMNDANKNISEVL 477

Query 326 VIDDIAEQTMMLALNASIEAARAGEGFAVVADEKVALAEESESTROSTVEELVEQMOP 385
 Database 478 VIGALSEQTNNLALNATEAARAGEGQRGAFAVVADEVRAAARTONSTLISDMILKLL 537

Query 386 ETEETVQDLEINQRIGEGVVEREEAMETLOBITDVEDAASGMQEVSSTATEQAVSTEE 445
 Database 538 GTDSVVGAMERTKQQCOTTAJKTSEYSGSLNMMSASVSDDLSTQIAATEQOSTVAAE 597

Query 446 VAEMVDGVDDRAGEIAAALDDIADATDQQTRTVEEYRETCK 487
 Database 598 LSRNMLSIREVLSLVSGRTVGATESLSHSNHELQVSK 639

RESULT 15
 US-10-282-122A-69564
 / Sequence 69564, Application US/10282122A
 / Publication No. US20040029129A1
 / GENERAL INFORMATION:
 / APPLICANT: Wang, Liangsu
 / APPLICANT: Zamudio, Carlos
 / APPLICANT: Malone, Cheryl
 / APPLICANT: Haselbeck, Robert
 / APPLICANT: Zyskind, Judith
 / APPLICANT: Wall, Daniel
 / APPLICANT: Trawick, John
 / APPLICANT: Carr, Grant
 / APPLICANT: Yamamoto, Robert
 / APPLICANT: Forsyth, R.
 / APPLICANT: Xu, H.
 / TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 / FILE REFERENCE: ELTRA-034A
 / CURRENT APPLICATION NUMBER: US/10/282,122A

CURRENT FILING DATE: 2003-02-20
 / PRIOR APPLICATION NUMBER: 60/191,078
 / PRIOR FILING DATE: 2000-03-21
 / PRIOR APPLICATION NUMBER: 60/206,848
 / PRIOR FILING DATE: 2000-05-23
 / PRIOR APPLICATION NUMBER: 60/207,727
 / PRIOR FILING DATE: 2000-05-26
 / PRIOR APPLICATION NUMBER: 60/230,335
 / PRIOR FILING DATE: 2000-09-06
 / PRIOR APPLICATION NUMBER: 60/230,347
 / PRIOR FILING DATE: 2000-09-09
 / PRIOR APPLICATION NUMBER: 60/242,578
 / PRIOR FILING DATE: 2000-10-23
 / PRIOR APPLICATION NUMBER: 60/253,625
 / PRIOR FILING DATE: 2000-11-27
 / PRIOR APPLICATION NUMBER: 60/257,931
 / PRIOR FILING DATE: 2000-12-22
 / PRIOR APPLICATION NUMBER: 60/267,636
 / PRIOR FILING DATE: 2000-12-22
 / PRIOR FILING DATE: 2001-02-09
 / PRIOR FILING DATE: 2001-02-16
 / PRIOR FILING DATE: 2001-02-16
 / Remaining Prior Application data removed - See File Wrapper or PALM.
 / NUMBER OF SEQ ID NOS: 78614
 / SEQ ID NO: 69564
 / LENGTH: 501
 / TYPE: PRT
 / ORGANISM: *Pseudomonas syringae*
 US-10-282-122A-69564

Query Match 13.1%; Best Local Similarity 27.4%; Pred. No. 7.3e-13; Length 501;
 Matches 86; Conservative 59; Mismatches 136; Indels 33; Gaps 5;

Query 176 KLLTFLDQQIAMDITYD-----SYAQRQLHEDIDSQELANAVATHVEAPLSSLEATSQD 228
 Database 184 RMLATTINQFSPDSRIDERSGKHVSLAQRFEQFLAQITGLVGDVYRDTR --- GLGELGHD 239

Query 229 VAERTUTMR -----ARTDDQDRMADVSRSEISS-VSASVEVASTADDVRTSED 277
 Database 240 LAKASCSTLEGAQHOLUSELARMTGAMRMGADMNDISGHVAQAVORAGASDOV --- 293

Query 278 AEALAOGEAAAADDALATMTDIDEATDGTVTAGVBQLGRRAADESYSTG 337
 Database 294 -----AHGRDQDVDRASEITTOLAARISSTDDETQALANQBQIGKVLDVIGSIAQTNL 348

Query 338 ALNALSIAARAGEAGEGFAVVADEYKALAAESREQSOTREELVEQMQAETEETVQDDEV 397
 Database 349 ALNRAEYEARAGEGQRGAFAVVADEYRNLQRTASSTEKIQTIIEDIQKGSRQAATMSDS 408

Query 398 NORIGGVERVEAMRTLQEITDAADEAASEMQESTATEQAVSTEEVAEMVDGVDDRA 457
 Database 409 LGVGRCVENQRASSLSRGEGHTQINGLIATTEQQATASREIADQLRSYQAI 468

Query 458 GEIAAALDDIADAT 471
 Database 469 EHTAANGVLISS 482

Search completed: August 10, 2004, 15:42:26
 Job time : 73.0412 secs

BANK

C;Accession: T44973	C;Name: McCandless, J.; Banda, P.; Alam, M.
R;Zhang, W.; Brooun, A.;	Proc. Natl. Acad. Sci. U.S.A. 93, 4659-4654, 1996
A;Title: Signal transduction in the archaeon Halobacterium salinarium is processed through	
A;Reference number: Z22804;	
A;Accession: T44973	
A;Status: preliminary; translated from GB/EMBL/DDBJ	
A;Molecule type: DNA	
A;Cross-references: EMBL:U75435; NID:91654418; PID:NAB17880_1; PID:91654419	
A;Experimental source: strain Flix15	
A;Note: the source is designated as Halobacterium salinarium	
C;Genetics:	
A;Gene: htpIII; htp3; htA	
C;Superfamily: Halobacterium salinarum transducer protein car	
C;Keywords: methylated amino acid; signal transduction	
F/F;219-472/Region: MCP signalling domain similarity	
Query Match Score 609.5; DB 2; Length 482;	
Best Local Similarity 37.2%; Pred. No. 6.2e-20;	
Matches 181; Conservative 70; Mismatches 39; Gaps 12;	
Qy 15 NGIDGHALADRIGDEAEIAWRLSTFGIDDDTMALAAEOPLFTEATDAL-VTDFYDHL 73	
Db 19 NTLDVFAFA -VDAGDAVVAW-----DQIAAU----LETAPEAIGVTIDGERLN 63	
Qy 74 SYERTODLFANSTKTYEQLKETQAEXYLGLGRGEYDTEYAAQRARIGKIHDLVGLGPDVY 133	
Db 64 D---DGSRALANKVADTPIDAHHEY---DSVGLADESAYALLTGDDYVVEDTTVAGNTDLW 116	
Qy 134 LGATGTRYTGYLLDALADDVADGEAAAVALDLVRFLPMLKLL----TEDQOI--- 184	
Db 117 FIATPVHTGEFRGV1-EIVQDR-SSSARYQSEIQLAFGELEYLVTDAYDAGRFDATVDIA 174	
Qy 185 AMDTYIDSQAORLHDEIDSROELANAVATHYPEAFLSSLEATSQDYAERTDTMRARDQV 244	
Db 175 ABDTLLEDEYQIGRNUTIEFGDLAHAHITENVNDVERLEASQAVSESSAEIDELSTAQS 234	
Qy 245 DRMADVREISSVVASVVEVASTADDYRTRTSEDAEAIAAQGEEAADDALATMTDIDEATD 304	
Db 235 TNVSTVATEVEETLSATQYQIASTADEVYDTSATAERLADDGSAASDAADMNAVATAAD 294	
Qy 305 GYTAGVBOLGEAADVSVTGVIDIAEQTMNLALNASTEARAGEAGEFNAVDEVKA 364	
Db 295 SVTSDEALQNRIEDIVEDEVDTGTIAEQTMNLALNASTEAPPGEGEFFGFAVVAERVK 354	
Qy 365 LAFESREOSTRVYELVEMQAEETTYQDLDENVNIGEVGRVEEAMETLQEITDAVE- 423	
Db 355 LADDAQSAGHIESLVSEIQRDPAITVDTLVTITDRDEVAQVEDAMASFEETVAVEA 414	
Qy 424 --DAAAGMQEVSTATDEQAVSTTEVAEMYDGVDDRGESIAAAIDDIADATDQOVRTVEEV 481	
Db 415 TABRASNMM--VSQATNEQAAASAREIAAMVDETLADDITTAVADIVSQTEAGSAMHLHD 472	
Qy 482 REVVGKL 488	
Db 473 DESVSEL 479	
RESULT 5	
HB4336 Ht-3 transducer [imported] - Halobacterium sp. NRC-1	
C;Species: Halobacterium sp. NRC-1	
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002	
C;Accession: HB4336	
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithauser, B.; Keller, K.; Cruz, R.; Dawson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabbung, K.H.; Alam, M.; Freitas, T.	
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000	
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebbhardt, H.; Lowe, T.M.; I	
A;Title: Genome sequence of Halobacterium species NRC-1.	
A;Reference number: A84160; PMID:11016550	
A;Accession: HB4336	
A;Reference number: A84160; PMID:20504493; PMID:11016550	

A;Note: the source is designated as Halobacterium salinarum	Db	286 MSDRAFAIAAGDIDTAIEETRIDEVG-----DLRRSFRDIQEYIQLQTVAQQA DALAEQD 339
C;Genetics:	Qy	56 LFEATA-----DALYDFYD---HIESYNTTQDFPANSTTVEOLKETQEYLLG 102
A;Function:	Db	340 -FDAIDLDSKVPGRLGESLETMHWLETAIADELAQETABQRKEAEQSRE-EAEALA- 396
C;Superfamily: Halobacterium salinarum transducer protein htrII	Qy	103 LGREGYDTEYAAQRARIGKTHDVGLGPDYVLGAYTRYTGLLADADDVADRGEEAAA 162
C;Keywords: methylated amino acid; signal transduction; transmembrane protein	Db	397 -----FALLSQAQDIRTEVHAAD-----GDLTQRLTETDDHESWA 432
Query Match	Qy	398 EYLGLGRGEYDTEYAAQRARIGKTHDVGLGPDYVLGAYTRYTGLLADADDVADRGEEAAA 162
Best Local Similarity 33.8%; Score 597; DB 2; Length 805;	Db	399 YLLGKPLMILKLTEDQIAMDYIDSYARLHDEIPSQELANAVA--TIVEAPL 219
Matches 150; Conservative 76; Mismatches 158; Indels 60; Gaps 7;	Qy	400 AVDELVARFLPMILKLTEDQIAMDYIDSYARLHDEIPSQELANAVA--TIVEAPL 219
Db	401 385 LDIDLPGAFASLSOMHTRILEALTDLDEAREDAEQTRDAEAAASRLNBLERAA 444	
Qy	402 42 IDDTMAALLAEQPLFEATAALDLYDFYDLESYERT----QDLFANSKTKTVEQLKETQA 97	
Db	403 445 -----EYSDMEAALA-----GDLTR-----RLDEVVDSEPM 471	
Qy	404 98 EEAAGAVIDBELVARFLPMILKLTEDQIAMDYIDSYARLHDEIPSQELANAVATHYEA 217	
Db	405 472 QDIAEAFNDMMG-----DVEATL-----AQVRSIADAVD-----RASTDYST 508	
Qy	406 218 PLSSEATSDQDVATERDTMRARTDQVDRMADSREIISVSYASVEEVASTADDYRRTSED 277	
Db	407 509 SAEIIRSASDQVSSEVQDISADADQQRDLGTGVDEVSLSAVVEIASADDAVETQNQ 568	
Qy	408 278 AEALAOQGEAAADDALATMTDIDEATDGTAGVYTAGVYBQLGERAADVSITGVYDIAEQTML 337	
Db	409 569 RATESERQELGDAVLEREATADSAVERVTALEBAVDAIGDVTGCVITDAEQTML 628	
Qy	410 338 ALNASTIPARAGFAGEGFAVVADEVKALAEFREQSTTVEELVQMQABETEETVDQUDDEV 397	
Db	411 629 ALNANTIAARAKSGDFGFAVVADEVKOLADEYKESATEIETLVDDIADVAUTVADMSEL 688	
Qy	412 398 NORIGEGVERVERAMETQETDAVEDAASCMOEVSTDAEQAVSTBEAEMVGDYVDR 457	
Db	413 689 GERVDASETTEAALAAALDDGQEVANGSYOSISDADTDEQASTEVVIMIDEVTDSL 748	
Qy	414 458 GEIAAAALDIADATDQOVRTVEEV 481	
Db	415 749 DRTATESQVSAAEEOAASvSEV 772	
RESULT 7	Db	B84238 Htr18 transducer [Imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1	Qy	R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
C;Accession: B84238	Db	Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	Qy	Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
C;Accession: F84237	Db	A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	Qy	A;Text: Genome sequence of Halobacterium species NRC-1.
C;Accession: F84237	Db	Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	Qy	A;Reference number: A84160; MUID:20504483; PMID:11016950
C;Accession: F84237	Db	A;Status: preliminary
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	Qy	A;Molecule type: DNA
C;Accession: F84237	Db	Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	Qy	A;Residues: 1-778 <STO>
C;Accession: F84237	Db	C;Cross-references: GB:AE004437; PIDN:910580376; PIDN:AAG19266.1; GSPDB:GN00138
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	Qy	C;Genetics:
C;Accession: F84237	Db	A;Gene: htr4
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo	Qy	C;Superfamily: Halobacterium salinarum transducer protein htrII
C;Accession: B84238	Db	Query Match 24.1%; Score 578; DB 2; Length 778;
A;Status: preliminary	Qy	Best Local Similarity 33.7%; Pred. No. 2.6e-18;
A;Molecule type: DNA	Db	Matches 145; Conservative 75; Mismatches 172; Indels 38; Gaps 6;
A;Residues: 1-790 <STO>	Qy	147 -----ALADDVADRGEEAAAADVLARFL-----PMKLKLTEDPQIQIAMD 187
A;Cross-references: GB:AE004437; PIDN:910580382; PIDN:AAG19270.1; GSPDB:GN00138	Db	379 AEAASEAAQAAQAAEAEAREESEAQARLETTAEAFSETMAYAACDPLTVRLD 438
C;Genetics:	Qy	188 TYDSYAQ-----RLHDEIDSQELANAVATHYEPPLSLEATSQDVAE 231
A;Gene: htr18	Db	439 ADVEQQAMDIAAAFNEMAAADMTEATIADYVVAFADEVATA-STDASDAAAVEQTGRDVS 497
C;Superfamily: Halobacterium salinarum transducer protein htrII	Qy	1 MSNDNDTLVATADVRNGIDGHALADRGLDEAEIAWRLSSFTGIDD--DTMIALA--ABQP 55

RESULT 9							
Qy T48897		transducer protein VI (hrp6, hrD) [imported] - Halobacterium salinarum					
C:Species: Halobacterium salinarum							
C:Date: 02-Jun-2000 #sequence_revision		02-Jun-2000 #text_change 02-Sep-2000					
C:Accession: T48897							
R;Rudolph, J.; Nordmann, B.; Storch, K.F.; Gruenberg, H.; Rodewald, K.; Oesterreich, FEMS Microbiol. Lett. 139, 161-168, 1996							
A;Title: A family of halobacterial transducer proteins.							
A;Reference number: 224094; MUID:96275896; PMID:8674984							
A;Accession: T48897							
A;Status: preliminary; translated from GB/EMBL/DDBJ							
A;Molecule type: DNA							
A;Cross-references: EMBL:X95590; NID:91435133; PIDN:CAA64843.1; PMID:91435134							
A;Experimental source: strain S6							
C:Genetics:							
C:Superfamily: Halobacterium salinarum transducer protein hrII							
Qy T48897		Query Match					
C:Species: Halobacterium salinarum		23.9% Score 573; DB 2; Length 778;					
C:Date: 02-Jun-2000 #sequence_revision		Best Local Similarity 33.0%; Pred. No. 4.2e-18;					
C:Accession: T48897		Matches 144; Conservative 74; Mismatches 168; Indels 50; Gaps 6					
Qy T48897		Qy 88 TVEQKETQAEYLGLGRGEYDTETAAQRA RIGKIHDLVGLGPDVYLGAATRYTYIIL 146					
Db		320 TVRANDEAK-AALERGEYDTDDLVARYDELVPLRDIYQARIIRDNEQOVD 378					
Qy T48897		Qy 147 ----- ALADDVADRGEEAAAVADELVARFL----- PMKLKLTFDQQIAMD 187					
Db		379 AEAARSEAEAQQADEAAQAEAAQEESEQAQRLETTAAFSFTMRAYAAGDLTVRLD 438					
Qy T48897		Qy 188 TYIDSQA ----- RLIHDEDSRQEIANAVATHVEAPLSSLEAT 225					
Db		439 ADVEQAMADIAAFAENMAADMATIADAGRFAD----- ASTDADSAAVEQT 491					
Qy T48897		Qy 226 SQDVAERTDTMRARTDDQVDRMADVSEIISVSYSEBEVASTADVERTSDEAELAQG 285					
Db		492 GRD2DASVAGTRDRAAQRDQLEAVASBTDENSATIEVAASADQVAETSQRRAALGDD 551					
Qy T48897		Qy 286 EAADDALATMDIDEATDGVTAGVEQLGERAADVESVTGVIDDIAEQTMNLALNASTEA 345					
Db		552 QAAQDAVAQLEEEDETQAAATAVDDLEAKNSIEITVAATIDIAEQTMNLALNATEA 611					
Qy T48897		Qy 346 ARAGEAGEGFVVVADEVKALAEESEREQSTTRVEEVLQMQAEETEVDLDEVNRIGEGY 405					
Db		612 ARADQDGDFAVVADKLADESKASAEEIAVVAEVRQAEVSVAAMDRIQRSVSDGY 671					
Qy T48897		Qy 406 ERVEEAMETLQLTDAEDAAAGMQEVTSTDEAOVSTBEAENVHDYDGRGEITAALD 465					
Db		672 ETVSETERSLSLEAGRIBADTGQEIISNAAMDQDAASVSDVTAVGDVVAALGETATEAE 733					

P84327	Htr5 transducer [imported] - Halobacterium sp. NRC-1	A;Experimental source: strain S9 C;Species: Halobacterium salinarum transducer protein htrII C;Accession: P84327 R;Jung, K.H.; Alam, M.; Freitas, T.; Leithauer, B.; Storch, K.F.; Gruenberg, H.; Rodewald, K.; Oesterhelt, D.	Qy 84 NSTKTYVEQLKETOAEYLIGRLGRGYDTEYAAQR-----ARIGKHDVGLGPDVYLG 135 Db 346 NTAAAVQSLSSAAA-----IEAGNYDVGASSRDEIGOLFASGSMRDLVTOIDEAEA 401 Qy 136 ATYYTYTGILDALAD-----DVVADEGEAAAVDELVA-----RFPLMLKL-----178 Db 402 AREQATEAQODDAEERERAEDARRERAEDAKADALAALEQAERYSDVMMAZADGDLT 461 Qy 179 -----TFDQQIA-MDYYIDSAYAQRHLDEIDSRQELANAVAT--HVEAPLSSLEATS 226 Db 462 RRMPADDTNEAMAAIAASFNEMLAQWHTIIDI-QEFADAVATASEAEVGADAERAS 520 Qy 227 QDVAERTDTMRARTDDQVDRMADSREISSVSASVVEYASTADDVVRTSDEAALAQGQE 286 Db 521 GQYSESVQEIAGADEQRMNLDTSVSGENTDLSAEEVAAASDVAEHSQTAIARDCE 580 Qy 287 AAADDALATMTDIDBTDATDGVTAGEQLGERAADVESVTGVIDDABQTNMLNALNASEAA 346 Db 581 QTADDAEIRSLSVQEADATVQNEALDQMAPSETIVLISIQAETQTNMLNALNASEAA 640 Qy 347 RAGEAGEFPAVVADEKALAEEREPQSTRVEELVQMOAETEETVDOLDEVNORIGEGVE 406 Db 641 RADKSGDFAVVADEVKDAAETQESADDIERITETQSQTATAVARAAEESMDAGID 700 Qy 407 RVEEAMETLQEITDVAEADASGMOEVSTATDEQAVSTEEVAEMDGVDDRAGRIAALDD 466 Db 701 AVEEVDAFTASDHADETDTGQEBISDTDDQASSTEAVSMTTEEVADLSSTDGAQAS 760 Qy 467 IADATDOQVRTEVEVRETVYGKLS 489 Db 761 VSAAAEQQASMSEISDSVESLS 783	RESULT 13 T44597 transducer protein htrVI [similarity] - Halobacterium salinarum N;Alternate names: methyl-1-accepting transducer protein htr6 C;Species: Halobacterium salinarum C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000 C;Accession: T44597 Proc. Natl. Acad. Sci. U.S.A. 93, 4649-4654, 1996 A;Title: Signal transduction in the archaeon Halobacterium salinarum is processed through MCP signalling domain similarity A;Reference number: Z22804; PMID:96209786; PMID:8643458 A;Accession: T44597 A;Status: preliminary; translated from GB/EMBL/DDJB A;Molecule type: DNA A;Residues: 1-777 <ZHA> A;Cross-references: EMBL:U75338; NID:91654424; PID:AAB17883.1; PID:91654425
T46810	Halobacterial transducer protein IV [imported] - Halobacterium salinarum	A;Experimental source: strain mFLx15 C;Species: Halobacterium salinarum C;Accession: T46810 R;Rudolph, J.; Nordmann, B.; Storch, K.F.; Gruenberg, H.; Rodewald, K.; Oesterhelt, D.	Qy 88 TVEQLEKETOAEYLIGRLGRGYDTEYAAQR-RIGKHDVGLGPDVYLGATRYTGYL 146 Db 320 TVRALNDLEAK-APLRLERGEYDIDLVAVRDQASRDLFAASLRLFEASLRLFAVQDSDANEQQVD 378	
RESULT 12				
T46810				
C;Species: Halobacterium salinarum				
C;Accession: T46810				
R;FEMS Microbiol. Lett. 119, 161-168, 1996				
A;Title: A family of halobacterial transducer proteins.				
A;Reference number: Z24094; PMID:96275895; PMID:8674394				
A;Molecule type: DNA				
A;Residues: 1-810 <RUD>				
A;Cross-references: EMBL:X95589; NID:91435130; PMID:CAA64841.1; PMID:91435131				

QY	147 ALA-----DDVYADRGEAAAVALDELVARFLP 173	249 DVSRRISSVSASVEAVASTADDYRRTSEDAEALQQGEAAADDALATMIDIDEATDGVTIA 308
Db	379 AEARSEAEAAQAEPEAAQAEAEAREESIDAGAPGDDRRGVGLGDDACYAAGDLTVR --- 435	260 QIADEMDVSVTTEVAASADDIAXTARAEEAEGFAGRETAETAITMNEVESRTEQAYA 319
QY	174 MLKLTDFDQIA---MDTYTDSYAQRLHBDIBSDRQEELANAVA---THVEAPLSSLEATS 226	309 GVEQLGERAADVESVTGVDDIAEOTNMALNASTEARAEAGEFGFAVVADEVKLAEE 368
Db	436 ---LDADYQQAAMADMRRAFNEMADMEATIAERGAFADEVATPASTDASDAAAVECIG 491	320 SMEELNDYREIGEVSEMIADIASBNTNLNASTEARADGNSEGFAVVADEVKLAEE 379
QY	227 QDVABERTDTMRARTDDQDVERMADVSVREISSVSASVEAVASTADDYRRTSDEAALQQGE 286	369 SREQSTRVVELVQMQAEIEETVDOLDEVNQRICSEGVERVEANEWTLOETITDAEDAASG 428
Db	492 RDVSTAVGRTRDRAADQDLEAVESETDMSATBEVARVQVAETSSDRAAALGDDGQ 551	380 TKAATEEIDLIGTYQDRCTTVDVDIRETSDQVSSEGVEETDVTDALELTVDSVRTNDG 439
QY	287 AAADDALATMIDIDEATDGUTAGVADSVTGVDDIAEOTNMALNASTEA 346	429 MQEVTADTEQAVSPEVAEMKVGDYDRAZELIAAALDDIADATDOQVRTVEEVRETVGKL 488
QY	552 AAAQDAAVQLEEFEDTQAAATAVDLEAKMSEIETTAATDIAEOTNMALNANTEAA 611	440 IQEINGSTDQAQADAQQKATIMVEDMATAEQTOSDATAETTETQAESEYKEVFOLIDGL 499
QY	347 RAGEAGEFRAVVADEYKALABEESREOSTRYELVQMEATEETVDQDLEVNQRIGEGVE 406	489 S 489
Db	612 RADQDGFGFVVAADEVKDLADESKALAAEMQALVAEVRAQETSTSAAAMDRIQERYSDGV 671	Db 500 S 500
QY	407 RVEEAMETLQETTDAEVAASGMQEVSTATDEQAVSTEEVYAEVMGDVYDDRAZEEIAALDD 466	RESULT 15
Db	672 TVSETERSLSLDAIGRTAEPDTGVQEISNMDQQAASVSDVTTAVGVDVAALGEETATEAS 731	A84252
QY	467 IADATDQVRTVEEV 481	Htr15 transducer [imported] - Halobacterium sp. NRC-1
Db	732 TRDAAAABQATLSDV 746	C;Species: Halobacterium sp. NRC-1 C;Accession: A84252 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
QY	55299 sensory rhodopsin II transducer protein - Natronobacterium pharaonis	A;Title: Genome sequence of Halobacterium species NRC-1. A;Reference number: A84160; PMID:20304483; PMID:11016950
Db	55299 R;Seidel, R.; Scharf, B.; Gautei, M.; Kleine, K.; Oesterhelt, D.; Engelhard, M.	A;Accession: A84252 A;Status: preliminary A;Molecule type: <ST0> A;Cross-references: GB:AE004437; NID:910580513; PIDN:AA119381.1; GSPDB:GN00138 C;Genetics:
QY	55299 Proc. Natl. Acad. Sci. U.S.A. 92, 3036-3040, 1995	A;Residues: 1-636 <ST0>
Db	55299 A;Accession: S55296; MUID:7724074; PMID:7708770	A;Gene: htr15
QY	55299 A;Molecule type: DNA A;Residues: 1-534 <SER>	Query Match Score 22.4% Best Local Similarity 29.7% Matches 149; Conservatlive 99; Mismatches 191; Indels 63; Gaps 11;
Db	55299 A;Cross-references: EMBL:25086; NID:9510868; PID:g510869 A;Experimental source: strain SPI	QY 17 IDGHALADRIGLDEATIARWLSPFTGIDDDTMALAAEOPLFET----ADALVTD---- 67 Db 165 VDDDALIDSIGMP -- VFMFDATG----SVVANNETIEELTGCSRADAMGMDHSEA 214
QY	55299 C;Note: it is uncertain whether Met-1 or Met-22 is the initiator	QY 68 FYDHLEESTERTQDIFANSSTKTVFOLKEAQAEYLIGLRGEYDTYAAQFARIKGHDVLG 127 Db 215 FYPDDRRTKLAD----KVIERSPSAA----EFDTERSEEKAOLYADTSYM- 258
Db	55299 C;Superfamily: Halobacterium salinarum transducer protein	QY 128 LGPDVYLGAATRYTGLDALLAD----VADRGEEA--AAVADFLVARFLPML- 175 Db 259 ----TDOGSTDRHTRFKASPIFDDGELIAVETIEDTREDVDRADAEBLDELSITID 314
QY	55299 F:22-42/Domain: transmembrane #status predicted <TM1>	QY 176 -----KLTTFDQIQAMDTYLSYAQLHDEISQBLANAVATHVEAPLSSLEATSQ 227 Db 315 ALSSQQLSKRASFHEGELINEQLVSVSALNGMDQFERLVQGYDGTQBLADTTERATA 374
Db	55299 F:59-79/Domain: transmembrane #status predicted <TM2>	QY 228 DVABDTDTMARTDDQDMVDESERIVSASVEVASTADDVRRTSEIAEAIAQQGA 287 Db 375 DATDADTV ---SSQNEMUSSAANEMENFSASMQEVASSDQVASAQBQDAEASGLE 430
QY	55299 Query Match Score 22.5%; Length 534; Best Local Similarity 31.2%; Pred. No. 9.7e-17; Matches 150; Conservatlive 95; Mismatches 177; Indels 59; Gaps 12;	QY 288 ADDALATMIDIDEATDGVTGAGBOLGERAADVESVTGVDDIAEQTNMALNASTEAR 347 Db 431 ASEGANQATNEVIDISDDLMESVSKLESRMDEIEVNVVTEIAEVADQTNLALNANTEAR 497
QY	16 GIDGHALADRIGLDEA---EIAWLSLFTGSDDDTMALAAEOPLFEEAADALYTFY-- 69	QY 177 -----EINTELQAEEARFGEYMDRCAD 175 Db 124 FDEMRSQVTSLEDAKNAREDEAQPKRAE-----EINTELQAEEARFGEYMDRCAD 175
Db	72 GINLGLVAAATLGDTAAKSRM-GDGDIDVEL-----ETREDEIGDLVIA 123	QY 129 GDPVYLGAATRYTGLDALLADVVADYRGEAAAVALDELVARFLPMLKLLFDQJAMDT 188 Db 176 GD-----FTQRLDAEDTN-----FAMQSEG---SENEM-MDG 204
QY	70 -DHLSEYRTQDILFANSTKTVFOLKEAQAEYLIGLRGEYDTEAAQRARIKGHDVLG 128	QY 178 -----FAMQSEG---SENEM-MDG 204 Db 205 -IEALYGRIERPADSEAAVRANAE---SVMEA-SEDVNRAVONISDAAGDQFETVQ 259
Db	124 FDEMRSQVTSLEDAKNAREDEAQPKRAE-----EINTELQAEEARFGEYMDRCAD 175	QY 348 AGEAGGGFAVVADEYKALABEESRESTRYELVEMQAEETEVTQDLEVNORIGEGVER 407
QY	189 YIDSYAQLHDEDSRQEELANAVATHVEAPLSSLEATSDQVAERTDTMMARTDDQDVIRMA 248	Db
Db	205 -IEALYGRIERPADSEAAVRANAE---SVMEA-SEDVNRAVONISDAAGDQFETVQ 259	

Db 491 AGAGGGFAVVADEVKELANETREHTERIAGSISDVQQANETVLAVEESEHQIHRAGDE 550
Qy 408 VEBAMETLQEITDAVEAASGNQEVSSTATDEAQAVSTEVARNVDGVDRAQEJIAALADDI 467
Db 551 IDDALTLEEATSVDEBAAATGTEVARANDEQASTVEDVIVTIEDYQQQABEAAAASDRI 610
Qy 468 ADATDOCVRTYEEVRETVGKL\$ 489
Db 611 VSATQBESTAVSQLSERVVKLT 632

Search completed: August 10, 2004, 15:29:26
Job time : 22.3123 secs

GenCore version 5.11.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: August 10, 2004, 15:19:00 ; Search time 15.3923 Seconds
 (without alignments)
 1654.229 Million cell updates/sec

title:	US-09-455-978B-2			
perfect score:	1 MSNDNTLVATDVRNGIDGH.....ATDQQVQRTVIEEVRETVGKLS 489			
sequence:				
oring table:	BLASTM62			
	Gapop 10.0 , Gapext 0.5			
searched:	141681 seqs, 52070155 residues			
total number of hits satisfying chosen parameters:	141681			
minimum DB seq length:	0			
maximum DB seq length:	200000000			
post-processing:	Minimum Match 0%	Match 100%	Maximum Match 100%	Listing First 45 summaries
database:	SwissProt_42;*			
Query	Score	Match	Length	ID
No.				
1	2394	100.0	489	1 HMAT HALSA
2	2360	98.6	489	1 HMAT HALN1
3	643	26.9	481	1 HTR3 HALN1
4	643	26.9	481	1 HTR3 HALSA
5	578	24.1	778	1 HTR6 HALN1
6	573	23.9	778	1 HTR6 HALSA
7	555	23.2	810	1 HTR4 HALN1
8	554	23.1	810	1 HALSA
9	537.5	22.5	534	1 HTR2 NATPH
10	521	21.8	545	1 HTR5 HALN1
11	492	20.6	433	1 HALVA
12	464	19.4	763	1 HTR2 HALN1
13	463.5	19.4	535	1 HTR1 HALN1
14	461	19.3	764	1 HALSA
15	449.5	18.8	535	1 HTR1 HALSA
16	412.5	17.2	432	1 HMAT BACSU
17	388.5	16.2	662	1 MCPB BACSU
18	386.5	16.1	682	1 PILJ PSEAE
19	379	15.8	662	1 TLPB BACSU
20	359.5	15.0	547	1 MCPG SALTY
21	359.5	15.0	1000	1 Y041 SYNY3
22	359.5	14.8	662	1 TLPB BACSU
23	351	14.7	548	1 HLYB VIBCH
24	349	14.6	553	1 MCP2 ECOLI
25	348.5	14.6	551	1 MCP1 ECOLI
26	347.5	14.5	573	1 TLPB BACSU
27	342.5	14.3	657	1 MCPA CAUCR
28	335.5	14.0	553	1 MCP2 SALTY
29	334	14.0	661	1 MCPA BACSU
30	333.5	13.9	654	1 MCPC BACSU
31	326.5	13.6	546	1 MCP3 ECOLI
32	323.5	13.5	668	1 MCP5 ENTHV
33	316.5	13.3	557	1 MCP6 DESH
34	316.5	13.3	669	1 MCP7 EENTAB
35	316.5	13.3	670	1 MCP8 ECAUOB
36	316.5	13.3	671	1 MCP9 ECAUOB
37	316.5	13.3	672	1 MCPA ECAUOB
38	316.5	13.3	673	1 MCPB ECAUOB
39	316.5	13.3	674	1 MCPD ECAUOB
40	316.5	13.3	675	1 MCPF ECAUOB
41	316.5	13.3	676	1 MCPG ECAUOB
42	316.5	13.3	677	1 MCPH ECAUOB
43	316.5	13.3	678	1 MCPJ ECAUOB
44	316.5	13.3	679	1 MCPK ECAUOB
45	316.5	13.3	680	1 MCPM ECAUOB

Description

P11413 halobacteri
 Q9P86 halobacteri
 Q9H20 halobacteri
 Q48316 halobacteri
 Q9B92 halobacteri
 Q48319 halobacteri
 Q9P84 halobacteri
 Q48317 halobacteri
 P42259 matronomona
 Q48318 halobacteri
 P42258 haloarcula
 Q9P81 halobacteri
 P33741 halobacteri
 P7140 salmonella
 P33935 halobacteri
 Q55445 synochocys
 P39216 bacillus sp
 P39215 bacillus sp
 P42257 pseudomonas
 P39217 bacillus sp
 P20275 salmonella
 P33926 bacillus sp
 P39214 bacillus sp
 P54576 bacillus sp
 P05704 escherichia
 P39206 bacillus sp
 Q00986 caulobacte
 P02941 salmonella
 P39214 bacillus sp
 P54576 bacillus sp
 P05704 escherichi
 P35847 escherichi
 P21822 enterobac

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

PT	DOMAIN	218 AA;	454	METHYL-ACCEPTING TRANSDUCER.				
SQ	SEQUENCE	489 AA;	52816 MW;	8457263 FCFCF616BFF CRC64;				
	Query Match	100.0%;	Score: 2394;	DB 1; Length 489;				
	Best Local Similarity	100.0%;	Pred. No. 1. 4e-94;					
	Matches	489;	Conservative 0;	Missmatches 0;	Indels 0;	Gaps 0;		
Qy		1	MSNDNTLVTDYRNGIDGHALADRIGDEAIAWRLSFTGIDDDTMALAAEQPLEAT	60				
Db		1	MSNDNTLVTDYRNGIDGHALADRIGDEAIAWRLSFTGIDDDTMALAAEQPLEAT	60				
Qy		61	ADALVTFYDILESYETTQDLFANSKTVTEQIKETOLEYLIGLGRGEYDTEYAAQRARG	120				
Db		61	ADALVTFYDILESYETTQDLFANSKTVTEQIKETOLEYLIGLGRGEYDTEYAAQRARG	120				
Qy		121	KIHDVLGLGPDVYLGYTRYTGLLLADDDVADRGEEAAAVDELVALPMLKLTLF	180				
Db		121	KIHDVLGLGPDVYLGYTRYTGLLLADDDVADRGEEAAAVDELVALPMLKLTLF	180				
Qy		181	DQQIAMDTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART	240				
Db		181	DQQIAMDTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART	240				
Qy		241	DDQVDRMADSREISSSSAEEVEYASTADDVRRTSEDAEAQQAEEAADDALATMTID	300				
Db		241	DDQVDRMADSREISSSSAEEVEYASTADDVRRTSEDAEAQQAEEAADDALATMTID	300				
Qy		301	EATDGTVTAQEOLGERADEVASVYVQDQVAD	360				
Db		301	EATDGTVTAQEOLGERADEVASVYVQDQVAD	360				
Qy		361	EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID	420				
Db		361	EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID	420				
Qy		421	AVEDAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE	480				
Db		421	AVEDAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE	480				
Qy		481	VRETVGKLS 489					
Db		481	VRETVGKLS 489					
	RESULT 2							
	HMAT HALN1	STANDARD;	PRT;	489 AA.				
AC	QSHPR5;							
DT	28-FEB-2003	(Rel. 41,	Created)					
DT	28-FEB-2003	(Rel. 41,	Last sequence update)					
DE		(Rel. 41,	Last annotation update)					
GN	HEMAT OR HTR10 OR VNG1505G.							
OS	Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).							
OC	Archaeal; Buryarchaeota; Halobacteria; Halobacteriales;							
OX	NCBI_TaxID=64091;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RA	Medline=2050483; PubMed=11016950;							
RA	NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,							
RA	Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbragna J.,							
RA	Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A., Leithauer B., Keller K., Cruz R., Dawson M.J., Hough D.W.,							
RA	Maddock D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Isenbarger T.A., Peck R.F., Pohlshoer M., Spindich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A., Leithauer B., Keller K., Cruz R., Dawson M.J., Hough D.W.,							
RA	Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).							
CC	-!- FUNCTION: Heme-containing signal transducer responsible for aerotaxis, the migratory response toward or away from oxygen (By similarity).							
CC								
	CC	-!- SUBUNIT: Homotetramer (Probable).						
	CC	-!- SIMILARITY: Contains 1 methyl-accepting transducer domain.						
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).						
	CC	EMBL; AE00564; AAC1.801.1; - .						
	DR	PIR; E84304; E84304.						
	DR	HSP; P02342; 1QUT.						
	DR	Intertpro; IPR04089; Chm1axis transduc.						
	DR	Interpro; IPR04090; Me chemotaxis.						
	DR	Pfam; PF00015; MCPSignal_1.						
	DR	PRINTS; PR0260; CHEMTRANSUDCR.						
	DR	SMART; SM00283; MA_1.						
	DR	PROSITE; PS50111; CHEMOTAXIS_TRANSUDC_2; 1.						
	KW	Transducer; Name; Complete proteo.						
	FT	DOMAIN 2360; DB 1; Length 489;						
	FT	Best Local Similarity 99.0%; Pred. No. 3.9e-93;						
	FT	Matches 494; Conservative 1; Mismatches 4; Indels 0; Gaps 0;						
	Qy	1 MSNDNTLVTDYRNGIDGHALADRIGDEAIAWRLSFTGIDDDTMALAAEQPLFEAT 60						
	Db	1 MSNDNTLVTDYRNGIDGHALADRIGDEAIAWRLSFTGIDDDTMALAAEQPLFEAT 60						
	Qy	2.8 DOMAIN 2360; DB 1; Length 489;						
	Qy	489 AA; 52852 MW; 802AEF179667791E CRC64;						
	Query Match	98.6%; Score 2360;						
	Match	Best Local Similarity 99.0%; Pred. No. 3.9e-93;						
	Matches	494; Conservative 1; Mismatches 4; Indels 0; Gaps 0;						
	Qy	1 MSNDNTLVTDYRNGIDGHALADRIGDEAIAWRLSFTGIDDDTMALAAEQPLFEAT 60						
	Db	1 MSNDNTLVTDYRNGIDGHALADRIGDEAIAWRLSFTGIDDDTMALAAEQPLFEAT 60						
	Qy	2.8 DOMAIN 2360; DB 1; Length 489;						
	Db	61 ADALVTFYDILESYETTQDLFANSKTVTEQIKETOLEYLIGLGRGEYDTEYAAQRARG 240						
	Qy	61 ADALVTFYDILESYETTQDLFANSKTVTEQIKETOLEYLIGLGRGEYDTEYAAQRARG 240						
	Db	61 ADALVTFYDILESYETTQDLFANSKTVTEQIKETOLEYLIGLGRGEYDTEYAAQRARG 240						
	Qy	121 KIHDVLGLGPDVYLGYTRYTGLLLADDDVADRGEEAAAVDELVALPMLKLTF 180						
	Db	61 KIHDVLGLGPDVYLGYTRYTGLLLADDDVADRGEEAAAVDELVALPMLKLTF 180						
	Qy	121 KIHDVLGLGPDVYLGYTRYTGLLLADDDVADRGEEAAAVDELVALPMLKLTF 180						
	Db	121 KIHDVLGLGPDVYLGYTRYTGLLLADDDVADRGEEAAAVDELVALPMLKLTF 180						
	Qy	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Db	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Qy	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Db	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Qy	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Db	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Qy	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Db	181 DQ1AMTYIDSYAORLHDIDSROBLANAVATHVEAQLSSLEATSQDAERTDTMART 360						
	Qy	361 EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID 420						
	Db	361 EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID 420						
	Qy	361 EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID 420						
	Db	361 EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID 420						
	Qy	361 EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID 420						
	Db	361 EVKLAEEFREQSTRVELVQMOAETTETDVLDEVNORIGEVVEREAMETLQETID 420						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Qy	421 AVEADAASSMQEVSTATDQAVSTEEVAEMVDGYDRAESEAIAALDDIADATDQVRTVE						
	Db	421 AVEADAASSMQEVSTAT						

Best Local Similarity 37.6%; Pred. No. 8.5e-21; Matches 182; Conservative 72; Mismatches 196; Indels 34; Gaps 10;

Qy 15 NGIDGHALADRIGLEDEARIWRLSFTGIDDMAAELAEQPLFEATADAL-YTDFYDHE 73
Db 19 NTLDVPAEA- -VDAVGAVAW - -DDQIAAL - -LETAPEAIGVTDIGERLN 63

Qy 74 SYERTQDIFANSTKTVKEQLKETQAELYLGLGRGEYDTEYYAQRARIGKHDVIGLGPDVY 133
Db 64 D---DGSRALANKVADTPIDARHEY-- -DGVLGLADEYAYLLTGDDYTYEDTTVAGNTDLW 116

Qy 134 LGAYTRYTYTGLLDALADYVADRGEEAAAVADLVARLPLMKLL---TFDQQI--- 184
Db 117 FLATPVHTVHGEFRGV1-BIVODR-SSSRYQSOELQALFGELVDTLDAYDAGRFDATDVA 174

Qy 165 AMDTYIDYEYAQRHLDEISROELANAVATHVEAPLSLSEATSDVAERTDMRARTDQV 244
Db 175 AEDTLDEIYIQIGRNLTGFDTLAAHITEVNDVERLEAAQAVSSESAEDELSTAQS 234

Qy 245 DRMADVSEISSASVYSEVASTADDYRTTSDEAELAQGEEAADALATMDIDEATD 304
Db 235 TNVSTVAEPEVETISATVIASTADEVVDTSAATERLDDGSAASADADMADAVATA 294

Qy 305 GVTAGVEOLGERAADVSVTGYDDIAFQTNMNLANASIEAARAGEAGEGFVVADEVKVA 364
Db 295 SVTSDVEALQNRIEDIDEVVDTGIAFQTNMNLANASIEAARAGEYEGFVVAEVKVA 354

Qy 365 LASESREGSTRVBELEVEMQAEPEETVQDLEVNQRIGEVVEREAMBTLOBIDTAVED 4.24
Db 355 LAEDAQSNAQHLSLVSIIQRDADTVTDLVTTDRIEDAVAQVEDAMASFEEIVTAVEA 4.14

Qy 425 AASGMQEYVSTATDEQAVSTEEVAAEMVGDVVDRAGEIAAADDIADATQQVRTVEYRET 4.84
Db 415 TARGIEQYSDATNEQAASEEIAAMVDETAIDADDITTAVIDVSQTEAQSTILXLDLES 4.74

Qy 485 VGRKL 488
Db 475 VSEL 478

RESULT 5
HTR6_HALN1 ID HTR6_HALN1 STANDARD PRT; 778 AA.
AC Q9H92; DT 16-OCT-2001 (Rel. 40, Created)
DE Halobacterial transducer protein VI.
GN HTR4 OR HTPV1 OR VNG0806G
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea: Euryarchaeota: Halobacteria: Halobacteriales;
OC Halobacteriaceae: Halobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RM MEDLINE=20514483; PubMed=11016950;
RN NG.W.V., Kennedy, S.P., Mahirras, G.C., Berquist, B., Pan, M., Shukla, H.D., Lasyk, S.R., Thorsson, V., Sbrogna, J., Swartz, S., Weir, D., Hall, J., Dahl, T.A., Welni, R., Goo, Y.A., Leithauser, B., Keller, K., Cruz, R., Danson, M.J., Hough, D.W., Maddocks, D.G., Jablonski, P.B., Krebs, M.P., Angevine, C.M., Dale, H., Isenbarger, T.A., Peck, R.F., Pohlschroder, M., Spudich, J.L., Jung, K.-H., Alam, M., Freitas, T., Hou, S., Daniels, C.J., Dennis, P.P., Omer, A.D., El-Habbard, H., Lowe, T.M., Liang, P., Riley, M., Hood, L., Dassarma, S.; Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).

CC -1- FUNCTION: Potentially involved in chemo- or phototactic signal transduction (By similarity);
CC -1- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC -1- SIMILARITY: Contains 2 HAMP domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC DR EMBL; AB005022; AAC19266..1; DR PIR; FB4237; FB4237.
CC DR HSPP; P0407; DR InterPro; IPR004089; Chmtaxis_transd.
CC DR InterPro; IPR003660; HAMP.
CC DR InterPro; IPR004090; Ne_chemotaxis.
CC DR Pfam; PF00672; HAMP; 1.
CC DR Pfam; PF00015; MCPsignall; 1.
CC DR PRINTS; PRO0060; CHMTRANSUDCR.
CC DR SMART; SM00304; HAMP; 2.
CC DR PROSITE; PS50111; CHEMOTAXIS_TRANSUDC_2; 1.
CC DR PROSITE; PS50185; HAMP; 2.
KW Transducer; Transmembrane; Complete proteome; Repeat.
FT DOMAIN 1 26 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 27 47 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 48 296 POTENTIAL.
FT DOMAIN 297 317 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 318 778 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 318 412 455 HAMP 2.
FT DOMAIN 484 720 METHYL-ACCEPTING TRANSDUCER.
SQ SEQUENCE 778 AA; 82076 MW; D7522-8727A4FA79 CRC64;
Query Match 24.1% Score 578; DB 1; Length 778;
Best Local Similarity 33.7% Pred. No. 7.e-18; Minmatches 75; Indels 38; Gaps 6;
Matches 145; Conservative 75; DB 379 AEARSSEAQAQDAAQAEAAQAEDESEAARRLETTAEAFSETRAYAAGDLTVRLD 4.38
Qy 88 TVEQLKETQATYLLGLRGERTYEAQRA-RIGKTHDVLGIGPDVLYGTRYTGILD 146
Db 320 TVALNDLEAK-AALERGEYDIDTDLVARYDVLRLPEAFASLRDTQARI RDNEQQVD 378
Qy 147 -----ALADVVADRGREAAAADVELVAFL-----PMKLKLTFDQQIAMD 187
Db 379 AEARSSEAQAQDAAQAEAAQAEDESEAARRLETTAEAFSETRAYAAGDLTVRLD 4.38
Qy 188 TYDLSYAQ-----RLHDEIDSQELANAVATHVEAPLSSLEATSDQVDE 231
Db 439 ADVEQANADIAAAFNENMADMEATIAVVAFADVEYDAAVETQGDRVD 4.97
Qy 232 RTDTMRARTDDQDMVSRTEISVSASVVEEVATADDVRETSDEAELAQGEAAD 291
Db 498 AVGRIRDRAAQDQLEAVASETDEMSEATIEVAASADQVATTSQRAAALGDGQAAQD 557
Qy 292 ALATM0D1DEATDGVTAGVEQLGERRADVESVQDIDIAEOTQMLNALNASIEAARAGEA 351
Db 558 AVAQLEETEDTOQAAATAVDLEAKMSLEIETIVATDIAEQTMMLNALNATEAARADQD 617
Qy 352 GEGFAYVADEKVALAESRSREQSTREVLEVNQOATEETVQDQLEVNQNRIGEVVEREEA 411
Db 618 GDGFAYVADEKVALADSKASAEEIAEVAEVTRQTSVAAMDRICOPRSVDSVETVSET 677
Qy 412 METLQETIDAVEDAASGMQEYVSTATDEQAVSTEEVAEMDGVDYDRAGEIAAALDDIADAT 471
Db 678 ERSLSSETAGRINAEADTGVOEISNMDQQAASVSDVTTAVGDVAALGEBETATFAESTADA 737
Qy 472 DQYRVTYBV 481
Db 738 AEQATLSDV 747
RESULT 6
HTR6_HALSA ID HTR6_HALSA STANDARD PRT; 778 AA.
AC Q48319; DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

FT DOMAIN	439	493	HAMP 2.
FT DOMAIN	512	748	METHYL.ACCEPTING TRANSDUCER.
SQ SEQUENCE	810 AA:	85219 MW;	4BF364B7D22BD80 CRC64;
Query Match	23.2%	Score 555;	DB 1; Length 810;
Best Local Similarity	32.1%	Pred. No. 7.	5.8E-17; Mismatches 174; Indels 42; Gaps 8;
Matches 142; Conservative 85; RTM 1.0			
Qy 84 NSTRKTVQELKETAAEYLGLGREGYDTEYAAQR----ARIGKHNDVGLGPDVYG 135			
Db 346 NTAAAVQOSLSAAAE---IEAGNYDVYVASSRDRDEICQLPASIGSMRDALVTQIDEAEA 401			
Qy 136 ATRYRYTQTLDALAD---DVAADRGEAAAANDELVA-----RFLPMLKLL-----178			
Db 402 ARQATEQQDATAAERPAEDAAERPAEDAAERPAEDAAERPAEDAAERPAEDAAACDSDIT 461			
Db 179 -----TFDOQIA----MDTYIDSYAORLHDEIDSROELANAYAT---HVEAPLSSLEPAT 226			
Db 462 RRMPPADDTDNEAMAAIASFNEMLAQWHTI-LDI-QEPAVATASEEEVGAAADEEAS 520			
Qy 227 QDVAERTUTMARTDDQVDRMADSRETSVSSASVSEEVASTADDYRTSSEDEALAAQGE 286			
Db 521 GOVSESVBIAAGADEQNMLDTVSGETNTDLSAEIEVTAASDVAEBSHOTAEIARDGE 580			
Qy 287 AADDALLMTMDIDEATGTVAGTCVQEQLGERAADVESVTGVIDDIABOTNMNLNASEAA 346			
Db 581 QTAEDATLRSLTQEATATVQNEAALDQMAESEYVDSLISIAEQTMNLNASEAA 640			
Qy 347 RAGBAGEGRFAVVADEVKLRAAESREQSTRVEEVQMOAETEETVQDDEVNQRIEGVE 406			
Db 641 RADKGSDGDPAVADEVKDLAETQESAGDIERITEVQSQTTAVAAZRAAEESMDGID 700			
Qy 407 RVEEAMETLOEITDAVEDAASGMQEVSTPATDEQAVSTEEVAAEVNGVDVDRAGEIAAALDD 466			
Db 701 AVEBVVDAFTAYSDHADDITDGQEISSDTDDOAASTEAVMSTEEVADLSDSTAGEQS 760			
Qy 467 IADATDQVRTVBVEVRTYVGKLS 489			
Db 761 VSAAAEEQAAASNEISIDESVLS 783			
RESULT 8			
HTR4_HALSA	STANDARD;	PRT;	810 AA.
ID HTR4_HALSA			
AC Q48317;			
DT 15-DEC-1998 (Rel. 37, Created)			
DT 15-DEC-1998 (Rel. 37, Last sequence update)			
DT 28-FEB-2003 (Rel. 41, Last annotation update)			
DE Halobacterium transducer protein IV.			
GN HTR5 OR HTPIV.			
OS Halobacterium salinarium.			
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
RL FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC FURCION: Potentially involved in chemo- or phototactic signal			
CC transduction.			
CC -1- SIMILARITY: Contains 1 methyl-accepting transducer domain.			
CC -1- SIMILARITY: Contains 2 HAMP domains.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC the European Bioinformatics Institute. There are no restrictions on its			
CC use by non-profit institutions as long as its content is in no way			
CC modified and this statement is not removed. Usage by and for commercial			
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC or send an email to license@isb-sib.ch).			
CC CC			
CC ENBL: X95589; CAD64841; 1;			
CC DR PIR: T46810; T46810.			
CC DR HSSP: P02942; 1Q07.			
CC DR InterPro: IPR004089; ChmTaxis_transd.			
CC DR InterPro: IPR003660; HAMP.			
CC Pfam: PF00672; HAMP; 1.			
CC Pfam: PF00187; MCPsignal; 1.			
CC SMART: SM00283; MA; 1.			
CC PROSITE: PS0085; HAMP; 2.			
CC Transducer; Transmembrane; Repeat.			
CC FT DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).			
CC FT DOMAIN 1 39 POTENTIAL.			
CC FT TRANSMEM 39 59 EXTRACELLULAR (POTENTIAL).			
CC FT TRANSMEM 323 POTENTIAL.			
CC FT DOMAIN 324 CYTOPLASMIC (POTENTIAL).			
CC FT DOMAIN 344 CYTOPLASMIC (POTENTIAL).			
CC FT DOMAIN 345 810 HAMP 1.			
CC FT DOMAIN 345 397 HAMP 2.			
CC FT DOMAIN 493 748 METHYL.ACCEPTING TRANSDUCER.			
CC SEQUENCE 810 AA; 85207 MW; 37B016046A39D9BA CRC64;			
Query Match	23.1%	Score 554;	DB 1; Length 810;
Best Local Similarity	32.1%	Pred. No. 8.3e-17;	
Matches 142; Conservative 85; Mismatches 174; Indels 42; Gaps 8;			
Qy 84 NSTKTVQELKETAAEYLGLGREGYDTEYAAQR----ARIGKHNDVGLGPDVYG 135			
Db 346 NTAAAVQSLSSAAAE---IEAGNYDVYVASSRDRDEICQLPASIGSMRDALVTQIDEAEA 401			
Qy 136 ATYRYTQTLDALAD---DVAADRGEAAAANDELVA-----RFLPMLKLL-----178			
Db 402 AREQATEQQDAEERERAERERADAREERADAKADEAALAALEQAEQARYSDVMACADGDLT 461			
Qy 179 -----TDFQQA----MDTYIDSYAQRHLDEDSRQELANAYAT--HVEAPLSSLRATS 226			
Db 462 RRMPADDUDNEAMAAIAASENMLAQMEHTLIDI-QFPAQDATASEEAAEVGAADAEFAS 520			
Qy 227 QDVAERTDTMRARTDDQVDRMADSRETSVSSAEEVADVRSTSEDEALAOQGE 286			
Db 521 GQYSESVQEIAGADEQRNRNMLDTSGEMTDLSSAIEEVAAASDVAEHSHQTAETARDGE 580			
Qy 287 AADDALATMTDIDEATGTVAGCEOLGERAADVESVTGVIDDIABOTNMNLNASEAA 346			
Db 581 QTMDAERSLVSQEA-DATVONVEALDDQMMEISBIVDLSIQAEQTMNLNASEAA 640			
Qy 347 RAGEAGEFAYVADKEVKALESREQSTREEVQMOAETEETVQDDEVNQRIEGVE 406			
Db 641 RADKGSDGDFAVADEQDVAEKTQESAGDIERITEVQDOLDEVNQRIGEGVE 406			
Qy 407 RVEEAMETLOEITDAVEDAASGMQEVSTPATDEQAVSTEEVAAEVNGVDVDRAGEIAAALDD 466			
Db 701 AVEBVVDAFTAYSDHADDITDGQEISSDTDDOAASTEAVMSTEEVADLSDSTAGEQS 760			
Qy 467 IADATDQVRTVBVEVRTYVGKLS 489			
Db 761 VSAAAEEQAAASNEISIDESVLS 783			
RESULTS 9			
HTR2_NATPH STANDARD; PRT; 534 AA.			
ID HTR2_NATPH			
AC P42259;			
DT 01-NOV-1995 (Rel. 32, Created)			
DT 01-NOV-1995 (Rel. 32, Last sequence update)			
CC 15-MAR-2004 (Rel. 43, Last annotation update)			
DB Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis			
DE protein III) (MPP-II).			
GN HTR2 OR HTRII.			
OS Natronomonas pharaonis (Natronobacterium pharaonis).			
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
RL FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			
CC FEMS Microbiol Lett 139:161-168(1996);			
CC MEDLINE=96275896; Published=8674984;			
RA Rudolph J., Nordmann B., Storch K.F., Gruenberg H., Rodewald K.,			
Oesterreit D.;			
"A family of halobacterial transducer proteins";			

OC	Halobacteriaceae; Natronomonas.
OX	[1] NCBI_TaxID=2257;
RN	SEQUENCE FROM N.A.
RP	STRAIN=SP-1 / 28;
RC	MEDLINE=92240470; PubMed=7708770;
RX	Seidel R., Scharf B., Gautel M., Kleine K., Oesterhelt D., Engelhard M.; "The primary structure of sensory rhodopsin II: a member of an additional retinal protein subgroup is coexpressed with its transducer, the halobacterial transducer of rhodopsin II."; Proc. Natl. Acad. Sci. U.S.A. 92:3036-3040(1995).
[2]	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 23-82. PMID=12369857; Gordejly V.I., Labahn J., Moukhamedzianov R., Efremov R., Granzin J., Schlesinger R., Bueldt G., Savopol T., Scheidig A.J., Klare J.P., Engelhard M.; "Molecular basis of transmembrane signalling by sensory rhodopsin II-transducin complex."; Nature 419:484-487(2002);
CC	-!- FUNCTION: Transduces signals from the phototaxis receptor sensory rhodopsin II (SR-II) to the flagellar motor. Responds to light changes through the variation of the level of methylation. Also acts as a chemotransducer.
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.
CC	-!- SIMILARITY: Contains 1 methyl accepting transducer domain.
CC	-!- SIMILARITY: Contains 2 HAMP domains.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	-----
DR	Z35086; CAA84468.1; -.
DR	PTR: S55299; S55299.
DR	PDB: 1H2S; 10-OCT-02.
DR	InterPro: IPR004089; Chmtaxis_transd.
DR	InterPro: IPR003660; HAMP.
DR	InterPro: IPR004090; Me_chemotaxis.
CC	-----
DR	Pfam: PF00672; HAMP; 1.
DR	Pfam: PF00015; MCPsignal; 1.
DR	PRINTS; PRO0260; CHEMTRNSDUCR.
DR	SMART; SM00304; HAMP; 2.
DR	SM00283; MA; 1.
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
KW	Transducer; Photoreceptor; Chemotaxis; Transmembrane; Methylation; Repeat; 3D-structure.
KW	DOMAIN 1 22 CYTOPLASMIC.
FT	TRANSMEM 23 43 TM1.
FT	DOMAIN 44 59 EXTRACELLULAR.
FT	TRANSMEM 60 81 TM2.
FT	DOMAIN 82 534 CYTOPLASMIC.
FT	DOMAIN 84 136 HAMP 1.
FT	DOMAIN 157 210 HAMP 2.
SQ	SEQUENCE 229 465 METHYL-ACCEPTING TRANSDUCER.
SQ	534 AA; 56622 MW; CBE8012C5F278E8 CRC64;
Query Match	22.5%; Score 537.5; DB 1; Length 534;
Best Local Similarity	31.2%; Pred No. 2.6e-16;
Matches	150+, Conservative 95%; Mismatches 177; Indels 59; Gaps 12;
Qy	16 GIDGHALADRIGLDEA--EIAWRLSFTGDDDTMAALLAEQPLFEATDALVDFY- -
Ds	72 GINLGIVVATIQTASLTAARKASRMG-DGDUDVEL----ETREDEIGDLYA 123
Qy	70 -DHLESVERTQDLFANSTKIVQEQLKETQAYEYLGLGRGEVTDYAAQRIGKIHDLVGL 128
Ds	124 FDEMRSQSVRTSLEDKRNAREDAEQQKRAE-----EINTELQREAERFEGVMRCA 175
Qy	129 GPDVYIGAYTRYTGTILDALADDVVADRGPEAAADEVFLPMLKLITFDQQIAMDT 188
Db	176 GD-----FTQRLDAFTDN-----SFNEM-MDG 204
Qy	189 YIDSYAQRLHEDIDSQELANAVATHVEAPISSELEATSQVAERTDTMRARTDDQVDRMA 248
Db	205 -IEALVGRIERFADAVSEDAEAVRAAE---SYNEA-SEDVNRAVONISDAAGDQTEVQ 259
Qy	249 DVSIETSSVASVERVYASTADDVRETTSEDAEALAOQGEAADALATMTDIDEATGVTA 308
Db	260 QIALEMDDVSATTEVAASADDIAKTRQDQLETGBAGRETAETATMEVNEVSRTQAVA 319
Qy	309 GVEQIGERAADVESVTGVDIADQETDIAETQNLAINSTAARAAAGEAGFGFVVADEVKLAFF 368
Db	320 SMEELNEDVREIGEVSEMTADAEONNIALNASEAARDGNSEGFAVVADEVKLAFF 379
Qy	369 SREQSTREVLEQVQMATEETVDOLDEVNRIGEGVERVEAMETLOEITDAVEDAASG 428
Db	380 TKAATPEEIDLIGTVQDRQTVDIRETSQSEGVEETDVTDALELVYDSTVETNDG 439
Qy	429 MQEVSTATDDEAVSTTEVAERKVGDGYDVRGEIAAALDIADATDQVTRVVEEVRETVGKL 488
Db	440 IQEINGSTDQAQADAQKATTMVEDMATSEQTASDAETETTQAESVKEVFDLIDGL 499
Qy	489 S 489
Db	500 S 500

RESULT 10
 HTR5_HALN1 STANDARD; PRT; 545 AA.
 ID HTR5_HALN1 MEDLINE=11016950;
 AC Q48318; Q9HP85;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Halobacterial transducer protein V.
 GN HTRP7 OR HTPV OR VNG1759G.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700932 / JCM 11081), and
 OS Halobacterium salinarium.
 OC Archaea; Buryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacteria; Halobacterium.
 NCBI TaxID=64091, 2242;
 RN [1] -
 RP SEQUENCE FROM N.A.
 STRAIN=NRC-1;
 RC MEDLINE=20504483; PubMed=11016950;
 RX RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., Shukla H.D., Lasky S.R., Thorsson N.S., Sbrrogna J., Swartzell S., Weir D., Keller K., Cruz R., Danson M.J., Hough D.W., Leithauer B., Isenbarger T.A., Peck R.F., Krebs C.M., Spudich J.L., Jung K.-H., Maddocks D.G., Jablonski P.E., Kohlschroder M., Angeline C.M., Dale H., Isenbarger T.A., Peck R.F., Kohlschroder M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ebhardt H., Love T.M., Liang P., Riley M., Hood L., DasSarma S.; RT "Genome sequence of Halobacterium species NRC-1"; RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000);
 RL [2] -
 RP SEQUENCE FROM N.A.
 RC SPECIES=H_salinari; STRAIN=S9;
 RX MEDLINE=92275896; PubMed=8674984;

RA Oesterhelt D.; Nordmann B.; Storch K.F., Gruenberg H., Rodewald K., Rudolph J.; Nordmann B.; Storch K.F., Gruenberg H., Rodewald K., Oesterhelt D.; "A family of halobacterial transducer proteins.";
 RT RL FEMS Microbiol. Lett. 139:161-168(1996).
 CC CC -!- FUNCTION: Potentially involved in chemo- or phototactic signal transduction.
 CC CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
 CC CC -!- SIMILARITY: Contains 1 HAMP domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - EMBL Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC EMBL; AE005080; AUG1995.1; -.
DR EMBL; 295589; CA64842.1; -.
DR PIR; B94327; B94327.
DR PIR; T6811; T46811.
DR B2942; 1007.
DR InterPro; IPR004889; Chmtaxis_transd.
DR InterPro; IPR003660; HAMP.
DR InterPro; IPR00490; Me chemotaxis.
DR Pfam; PF00015; MCPsignaT_1.
DR PRINTS; PR00260; CHEMTRNSD0CR.
DR SMART; SM00283; NA; 1.
DR SMART; SM00304; HAMP; 1.
DR PROSITE; PS00111; CHEMOTAXIS_TRANSUDC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
KW Transducer; Transmembrane; Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 80 100 POTENTIAL.
FT DOMAIN 164 217 HAMP.
FT DOMAIN 236 476 METHYL.ACCEPTING TRANSDUCER.
SEQUENCE 545 AA; 57070 MW; FDD870389C2F428B CRC64;
```

Query Match Score 51; DB 1; Length 545;
Best Local Similarity 29.6%; Pred. No. 1. 3e-15;
Matches 142; Conservative 94; Mismatches 195; Indels 48; Gaps 10;

```
28 LDEAETIAWRLSFTGIDDTMMAALAAEAPLFEATADALYTFD---YDHLESYERTQTLFA 83
Db 64 LEAGVATQI-FNLTGFLIMATTVAFFIGGTTATLA|VBDMEALVEBAQRQEAEERA 122
Qy 84 NSTKTVQELKETQAELYLLGLGRGEYDTYYAQPARIGKIHDLGIGPPVYIGAYTRYTG 143
Db 123 EAARRAEAEQKORE---AEROETAEEASAKOPAR-EFSAEYEQLAAD--DESQATEVGA 175
Qy 144 LLDALAD-----DVVADRGE--EAAADEVDELVARFLPMLKULTEQDQIAMDYIDSYA 194
Db 176 TLEAAASDSDLTARYDATTDNAEIAEVAVVNDM-----LTTMERTI----- 216
Qy 195 QRLHDEIDSROELANAVATHVEAPLSSLEATSDQVAERTDTMRARTDDQVRMDAYSREI 254
Db 217 ---DEIQGFSNTNTTASREATGAKAE1QDASATVSBSVQE1AAGTDDQRBQLESVAEEM 272
Qy 255 SSVASVSBEVASTADDVRTSEDAEAALAQOGEAAADDALATMDIDEATDGVTAGVEQLG 314
Db 273 DS-SATVSEBVAQSDVATAATTDVAGTKOTAEDIDDAQVQTMTQVANVDALE 332
Qy 315 ERAADVESVTGVNDDIAQTOQMMLALANASIEAARAGEA---GEGFAYVADEYKALABESR 370
Db 333 DLTEIDIAELSDIAQTMNLALANANIEAARAGSGCGNSNDGFAYVADEYKELATESQ 392
Qy 371 EQSTRVEELVQNQAEETBTVDLQDVEVNRIGGVERVEAMETLQEITDAVEDAASGMQ 430
Db 393 RSAKDIASLIEEVQSQNTTVEBIRVAQRVNDGAAWEETDAGAVENIQETDGVQ 452
Qy 431 EVSTATDFQAVSTEEVAVEMVGDDRAGEIAAALDDIAQDQVRYVEYVRETGVKL 489
Db 453 EISQAMDQAQRSERVVSVDIATISQDATAENVAASBQTASTEVSLSQSLA 511
Qy 511 IAMDTYIDSYAOBLHEDIDSROELANAVATHVEAPLSSLEATSDQVAERTDTMRARTDDQ 243
Db 512 1AMDTYIDSYAOBLHEDIDSROELANAVATHVEAPLSSLEATSDQVAERTDTMRARTDDQ 243
Qy 513 YDHLAESERTQDLFANSSTKTVEQLKETOAEYLLGIGREYDTEYAAQRARIGKIHDLVGL 128
Db 514 FDNMRANLRQTISEAETAKQEAKEAQQ-----AAAREDVESENEMEALTG----- 58
Qy 515 GPDVYLGAYTRYTGILDLAADF----VVADRGEAAAAYDELVARFLPMKLUITDQO 183
Db 516 -----HLELKAQOYSDAIDAANGDLTARVKTDMSMDAMAEGDINTLDALEDTADMK 114
Qy 517 1AMDTYIDSYAOBLHEDIDSROELANAVATHVEAPLSSLEATSDQVAERTDTMRARTDDQ 243
Db 518 1AMDTYIDSYAOBLHEDIDSROELANAVATHVEAPLSSLEATSDQVAERTDTMRARTDDQ 243
Qy 519 -----SDRVNSNAE-----RVDASKVSKSINEIFEGTEQ 155
Db 520 244 VDRMADVSRTEISSVASVEEYASTADDVRTSEDAEAALAQOGEAAADDALATMDIDBAT 303
Qy 521 156 NEGLESAAEMQNLISATAQVASSAQVADTSQAAKVGDGREAAQEAIAEMSIAEAT 215
Db 522 304 DGVTAQVEQLGRERADEVSYTSEVIIIDDAEOTNMMLANASIEAARAGEAGEFPAVADBEK 363.
Db 523 216 GETVEEINALDDELDGEIGEVGTVSIVEQTMMLANASIEAHDGDGEFPAVADBEIK 275
DE Sensorzy rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis
```

protein II) (MPP-II) (Fragment).

HTRII.

Haloarcula vallis-mortis.

Archaea; Euryarchaeota; Halobacteria; Halobacteriales;

Halobacteriaceae; Haloarcula.

NCEP_ID=28442;

RN [1]

SEQUENCE FROM N A.

STRAINNAME 29715;

RX MEDLINE=95224074; PubMed=7708770;

RX Seidel R., Scharf B., Gauteil M., Kleine K., Oesterheld D.,

Engelhardt M.,

RT "The primary structure of sensory rhodopsin II: a member of an additional retinal protein subgroup is coexpressed with its

RT transducin, the halobacterial transducin of rhodopsin II."

RT Proc. Natl. Acad. Sci. U.S.A. 92:3036-3040(1995).

CC -!- FUNCTION: Transduces signals from the phototaxis receptor sensory

CC rhodopsin II (SR-II) to the flagellar motor. Responds to light

CC changes through the variation of the level of methylation. Also

CC acts as a chemotransducer.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).

CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.

CC -!- SIMILARITY: Contains 1 HAMP domain.

CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC DR EMBL; Z35308; CA84549; 1; -.

CC DR HSSP; P02942; 1Q07.

CC DR InterPro; IPR004089; Chmtaxis_transd.

CC DR InterPro; IPR003660; HAMP.

CC DR InterPro; IPR004090; Me chemotaxis.

CC DR PRINTS; PR00260; CHEMTRNSD0CR.

CC DR SMART; SM003204; HAMP; 1.

CC DR SMART; SM00283; MA; 1.

CC DR PROSITE; PSS0111; CHMOTAXIS_TRANSUDC_2; 1.

CC DR PROSITE; PSS0085; HAMP; 1.

CC KW Transducer; Photoreceptor; Chemotaxis; Transmembrane; Methylation.

CC FT NON_TER 1 1

CC FT DOMAIN 58 111 HAMP.

CC FT DOMAIN 130 366 METYL_ACCEPTING_TRANSDUCER.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86; GapS 5;

CC Matches 127; Conservative 86; DR SMART; SM00283; MA; 1.

CC SQ SEQUENCE 433 AA; 45935 MW; 90507BB897D943CO CRC64;

CC Query Match Score 492; DB 1; Length 433;

CC Best Local Similarity 29.9%; Pred. No. 1. 8e-14; Mismatches 86;

RN	[2] SEQUENCE FROM N.A.; AND SEQUENCE OF 1-10; 350-377 AND 457-476.	Db	169 ---LLTQRVDDTDHEAMETVGTAFNQNMDDLOATVTTVADEIAKTERMSSETSAIDI 224
RC	RP SPECIES:H.alobac.; STRAIN:FLXK;	Qy	230 -AERITDINRA---RTDDQVDRMDSREISSVSASVEEVASTADDVPRPSSEDAFALIA 282
RC	MEDLINE=93101637; PubMed=1465418;	Db	225 EASAGDTEAVSKIEQSANDORTELDSAADDYQQVSSAEEATAIIDDLASRSBDEVATAS 284
RA	Yao V.J., Spudich J.L.;	Qy	283 QQEAAADDALATMDIDEATGVTAGVEGLERAADVSTGVIDDIAEQTMNLANAS 342
RT	"Primary structure of an archaeabacterial transducer, a methyl-accepting protein associated with sensory rhodopsin I.";	Db	285 DAARDSSKSALDEMSSIETEVDAVQVEQLDQVATEIDIVDVTIGEQTMNLANAS 344
RT	Proc. Natl. Acad. Sci. U.S.A. 89:11915-11919 (1992).	CC	-1- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR
CC	SENSORY RHODOPSIN I (SR-I) TO THE FLAGELLAR MOTOR. RESPONDS TO LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.	CC	-1- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC	-1- SIMILARITY: Contains 2 HAMP domains.	CC	-1- SIMILARITY: Contains 1 methyl-accepting transducer domain.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to licensee@isb-sib.ch).	CC	-1- SIMILARITY: Contains 2 HAMP domains.
CC	EMBL; AE005075; AAG19913.1; -.	DR	401 IGFGRVERVEEAMETLQELTDAVEDAASGMQEVSTATDEQ-----AVSTEVFAEMYD-- 451
CC	EMBL; T05603; AA672315.1; -.	DR	405 VBSGSETVESTLDRIRTIADSTAEVNSIDEIQRTRTSEQAQETVQSTATSVRAGLSDDT 464
CC	PIR; AA4318; E84318.	DR	452 -----GVDDRAGEIAALDDIADATDQVQT-----VEEVRETGV 486
CC	InterPro; IPR02942; I0077	DB	465 TALASDASASAVIGORESSEETIASLEQFQNTAVEQLSRVASFTVATEDSETAG 518
CC	HSSP; P02942; I0077	Qy	RESULT 14
CC	InterPro; IPR004089; Chmtaxis_transd.	Db	HTR2_HALSA
CC	InterPro; IPR003660; HAMP.	AC	P71470;
CC	InterPro; IPR00490; Me_chemosensis.	DT	15-JUL-1998 (Rel. 36, Created)
CC	Pfam; PF0672; HAMP; 1.	DR	15-JUL-1998 (Rel. 36, Last sequence update)
CC	Pfam; PF00015; MCPsignal; 1.	DE	15-MAR-2004 (Rel. 43, Last annotation update)
CC	PRINTS; PR00260; CHEMTRNSNDICR.	DE	Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis protein II) (MPP-II).
CC	SMART; SM00304; HAMP; 2.	GN	HTR2 OR HTRII.
CC	SMART; SM00283; MA; 1.	OS	Halobacterium salinarium.
CC	PROSITE; PS50111; CHENOTAXIS_TRANSDUC_2; 1.	OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
CC	PROSITE; PSS0085; HAMP; 2.	OC	Halobacteriaceae; Halobacterium.
CC	Transducer; photoreceptor; transmembrane; Methylation; KW	DBP	NCBI_TaxID=2242;
CC	Complete proteome; Repeat; KW	RN	OX
CC	INIT_MET	RP	SEQUENCE FROM N.A.
FT	0	RC	STRAIN=FLX15;
FT	DOMAIN 1 13	RX	MEDLINE=96323203; PubMed=8710852;
FT	TRANSMEM 14 28	RA	Zhang W., Brooun A., Mueller M.M., Alam M.;
FT	DOMAIN 29 38	RA	"The primary structures of the Archaeon Halobacterium salinarum blue light receptor sensory rhodopsin II and its transducer, a methyl-
FT	TRANSMEM 39 54	RT	-1- accepting protein.";
FT	DOMAIN 55 535	RT	RT
FT	54 106	RL	Proc. Natl. Acad. Sci. U.S.A. 93:8230-8235 (1996).
FT	INIT_MET 0 0	RN	[1]
FT	DOMAIN 13 13	RP	FUNCTION.
FT	TRANSMEM 14 28	RX	PubMed=9515936;
FT	DOMAIN 29 38	RA	MEDLINE=96323203; PubMed=8710852;
FT	TRANSMEM 39 54	RT	"Sensory rhodopsin II transducer HTRII is also responsible for serine
FT	DOMAIN 55 535	RT	methylation in the archaeon Halobacterium salinarum.";
FT	54 106	RL	J. Bacteriol. 180:1600-1602 (1998).
FT	INIT_MET 0 0	CC	-1- FUNCTION: Transduces signals from the phototaxis receptor sensory rhodopsin II (SR-II) to the flagellar motor. Responds to light changes through the variation of the level of methylation. Also acts as a chemotransducer.
FT	DOMAIN 148 201	CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
FT	TRANSMEM 220 458	CC	-1- SIMILARITY: Contains 1 methyl-accepting transducer domain.
FT	MOD_RES 265 265	CC	-1- SIMILARITY: Contains 2 HAMP domains.
FT	MOD_RES 272 272	CC	-
FT	MOD_RES 279 279	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT	MOD_RES 463 463	CC	the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
FT	472 472	CC	-
FT	SEQUENCE 535 AA; 56544 MW; B99454F66A9D091. CRC64;	CC	-
CC	Query Match 19.4%; Score 463.5; DB 1; Length 535;	CC	-
CC	Best Local Similarity 27.5%; Pred. No. 3..e-13;	CC	-
CC	Matches 147; Conservative 87; Mismatches 205; Indels 95; Gaps 15;	CC	-
Qy	2 SNDNDTLYTADYVRNGIDGHALADRIGLEDAEIARWLSTGTDDDTMAlAAEOPLFETA 61	CC	-
Db	31 TNDVPSVIA---GIAG---LITLGSIINAAEVASIKEIAAQTERVANGNLQEBTSTRT 84	CC	-
Qy	62 DALVTDFTDHLEYERTQDFLNSTKTQ---LKEToAEYLGLGEYDTEYAAORA 117	CC	-
Db	85 D---ERGSLADTEQMOSLRGLNEMERTRADBEQAE---AEFAREEAEQARKOA 136	CC	-
Qy	118 RIGKHDVIGLGPDVYLGAYTRYTGLDALADDVARGRE-EAAAAMDVELARFLML 175	CC	-
Db	137 QAAERE-----ARELAATYQDTAKYGETMEAATGD----- 168	DR	EMBL; U62676; AAC44369.1; -.
Qy	176 KLLTFDQSQIAMDT----YIDSYAQRHDEIORSQELANAVATHVPLSSLEATSODV 229	DR	PTB; T44946; T44946.
Qy	177 IP004089; Chmtaxis_transd.	DR	HSSP; P02942; IQU7.
Qy	InterPro; IP004089; Chmtaxis_transd.	DR	InterPro; IP004089; Chmtaxis_transd.

DR InterPro; IPR003660; HAMP.
 DR PF000672; HAMP; 1.
 DR Pfam; PF00015; MCpsignal; 1.
 DR SMART; S00304; HAMP; 2.
 DR SMART; S000283; MA; 1.
 DR PROSITE; PS0111; CHENOTAXIS_TRANSDUC_2; 1.
 DR PROSITE; PSS0885; HAMP; 2.
 DR Transduc.; Photoreceptor; Chemotaxis; Transmembrane; Methylation;
 KW Repeat.
 FT INIT-MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 16 36 POTENTIAL.
 FT DOMAIN 37 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 298 POTENTIAL.
 FT DOMAIN 299 764 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 303 355 HAMP 1.
 FT DOMAIN 397 450 HAMP 2.
 FT DOMAIN 469 705 METHYL-ACCEPTING TRANSDUCER.
 SQ SEQUENCE 764 AA; 79187 MW; 1E0DB4E46FC588 CRC64;

Query Match 19.3%; Score 461; DB 1; Length 764;
 Best Local Similarity 28.7%; Pred. No. 6.4e-13;
 Matches 135; Conservative 75; Mismatches 219; Indels 42; Gaps 9;

Qy 37 LSFTGIDDDTMALAAEAPPLFEATADALVTDYDLYSERTQDLFANSTKTVEQLKETQ 96
 Db 294 LSLIGIVGSGTWTALRQ-FSRRADEAAGIDDTDSR-NDEFGTLLAEFRSMDLS 350

Qy 97 AEYLLGLGRGEYDTEYAAQRARIKGKHIDVLGPGDVLGAYTRYTGLLDAIADDVADIR 156
 Db 351 SESLT----DAERAATRAEDAR-EDAEQQRADA-----EAAREDAAAR 389

Qy 157 -GEAAAAAVDELVARFLPMLKLIT--FDQOIAMDTYDIDSQAQLRH----DEIDSQ 205
 Db 390 KDAQETARALESAADYEELATAVADGDLTRVDASTDHDAMARIHALNDMDDIETSV 449

Qy 206 ELANAVATHVEAPLSSLEA-----TSQVAERTDTMRARTDDQYDRMADVREISSVS 258
 Db 450 AATATFSDHVSDAAQVRVEDAAGDAGTAGTDTVADEVISDGGATEQTRLHVAGEYDLS 509

Qy 259 ASVEEVASTADDVVRTSEDAEAQQAEEAADDALATMDDEATGCVTAQYBQLGERAA 318
 Db 510 ASAAEVETAVASLADTAGQQAASAVDDDRQTADEAVIMDVAADDAAAADAMDALDSEMA 569

Qy 319 DVESTVGIVDIAEQTINMLAINASIEARAGAGEGFAVVADEKVALAEESREQSTRVE 378
 Db 570 DIGEIVDVIAADTADOTINMLAINASIEARTGADGDFAVADEVKTIAESDAADDIES 629

Qy 379 LVEQMQAETTVDOLDEVNORIGEVYVEREFAMEETLOITDAEQAAGMNEQESTATDE 438
 Db 630 RLLALQGQVSVDADEMATSDFVSDGRTAVGDAATADDDVSVFADIDAAQGQIRATOR 689

Qy 439 QAVSTEEVAEMYDGVDDRAGEETAAADDIAADTQQVRTVEEVRETIVGKS 489
 Db 690 QAHAAASRVASAVDEVAGISQETAAQATAVADSAAATQDTLSSVDDAAADDLA 740

RESULT 15
 HTR1_HALSA ID_HTR1_HALSA STANDARD; PRT; 535 AA.
 AC P3955; HTR1_HALSA STANDARD; PRT; 535 AA.
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Sensory rhodopsin I transducer (HTR1) (Methyl-accepting phototaxis protein I) (MPP-I).
 DE HTR1 OR HTR1 OR HTR.
 OS Halobacterium salinarium
 Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC NCBI_TaxID-2242;
 RN [1] SEQUENCE FROM N.A.
 RP

RC STRAIN=L33; RX MEDLINE=90384855; PubMed=2205844;
 RA Ferrando E., Krab M., Marwan W., Oesterhelt D.; Oesterhelt D.;
 RT "A gene from *S. pombe* with homology to *E. coli* RNase III blocks
 conjugation and sporulation when overexpressed in wild type cells."
 RL Nucleic Acids Res. 18:5304-5304 (1990).
 RN [2]
 RP FUNCTION.
 RX MEDLINE=94244615; PubMed=8187768;
 RA Krab M., Marwan W., Vermeijio A., Oesterhelt D.;
 RT "Phototaxis of Halobacterium salinarium requires a signalling complex
 of sensory rhodopsin I and its methyl-accepting transducer HTR1.";
 RL EMBO J. 13:2150-2155 (1994).
 CC -!- FUNCTION: TRANSDUCES SIGNALS FROM THE PHOTOTAXIS RECEPTOR.
 CC -!- SENSOR: RHODOPSIN I (SR-I) TO THE FLAGELLAR MOTOR. RESPONDS TO
 CC LIGHT CHANGES THROUGH THE VARIATION OF THE LEVEL OF METHYLATION.
 CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
 CC -!- SIMILARITY: Contains 2 HAMP domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC DR EMBL; X68591; CRA48578.1; -.
 CC DR HSSP; P02942; 1QUT; -.
 CC DR InterPro; IPR004089; Chmataxis_transd.
 CC DR InterPro; IPR004660; HAMP.
 CC DR InterPro; IPR004090; Me_chemotaxis.
 CC DR Pfam; PF00672; HAMP; 1.
 CC DR Pfam; PF00015; MCPsignal; 1.
 CC DR SMART; SM00304; HAMP; 2.
 CC DR SMART; SM0283; MA; 1.
 CC DR PROSITE; PS50011; CHEMOTAXIS_TRANSDUC_2; 1.
 CC DR PROSITE; PS50085; HAMP; 2.
 CC KW Transduc.; Photoreceptor; Transmembrane; Methylation; Repeat.
 CC FT INIT-MET 0 0 BY SIMILARITY.
 CC FT DOMAIN 1 8 CYTOPASIC (POTENTIAL).
 CC FT TRANSMEM 9 29 POTENTIAL.
 CC FT DOMAIN 30 33 EXTRACELLULAR (POTENTIAL).
 CC FT TRANSMEM 34 54 POTENTIAL.
 CC FT DOMAIN 55 535 CYTOPASIC (POTENTIAL).
 CC FT DOMAIN 54 106 HAMP 2.
 CC FT DOMAIN 148 201 HAMP 2.
 CC FT DOMAIN 220 458 METHYL-ACCEPTING TRANSDUCER.
 CC FT MOD_RES 265 272 METHYLATION (BY SIMILARITY).
 CC FT MOD_RES 279 279 METHYLATION (BY SIMILARITY).
 CC FT MOD_RES 463 463 METHYLATION (BY SIMILARITY).
 CC FT MOD_RES 472 472 METHYLATION (BY SIMILARITY).
 CC SQ SEQUENCE 535 AA; 56814 MW; 504A165AA7FA8A45 CRC64;

Query Match 18.8%; Score 449.5; DB 1; Length 535;
 Best Local Similarity 27.7%; Pred. No. 1.4e-12;
 Matches 141; Conservative 86; Mismatches 191; Indexes 91; Gaps 15;

Qy 28 LDEAEATWRLSFTGIDDDTMALAAEAPPLFEATADALVTDYDLYSERTQDLFANSTK 87
 Db 51 INAGETVGRKEIGACTERTANGUQEQTSTRD---EFGSGLADSIEMQSRGRIN 106

Qy 88 TVEQ---LKETOAFYLLGLGRGEYDTEYAQRARIKGKTHDVLGJGPDVYLGAATRYTG 143
 Db 107 EMERTSDLEETQTE---SETTREESEQKQEAOSAE-----REARE 145

Qy 144 LLDAADDVVAADRGE-EAAAANDVELVARLEPLMPKLTPDQI-----AMDIXIDSYA 194
 Db 146 LAATYQDSSNRD-GTTMESSATGQD-----FTQRVDVTDPEAMEFTGTAFN 190

Qy 195 QRLHDEIDSQELANAVATHVEAPLSSLTEATSQDV-AERTDTMRA-----RTDQVDRM 247

```

Db      191 QMM-DDQATVRTVVADEIITAKTERMSETSADIERASAGDTVEASKIESOPNDQRTEL 249
Qy      248 ADVSRELSVSASVEEVASTADDVRRTSEDABALAOQGEAAADDALATMDIDEATDGV 307
Db      250 DSAADDVQVSASAEETATIDDLASRSEDVATASDARDSSKSALDBMSSLETEVIDDPV 309
Qy      308 ACFYBQLGERAADVESVTYCIVDIAECQTMMLANASIRPARAG-EAGEGFAVYADEVKAL 365
Db      310 GOYEHRLRQVAELTDIVAVITDGEQTMMLANASIRPARAGNADCGFSVDADEVKDL 369
Qy      366 AEEFSREQTRVEELVNQAEETEVDOLDEVNQRIIGGVERVEAMBTLOEITDAVEDA 425
Db      370 AEETQDRANEIAAVVEKUTAQEDVTASIQHTRTRVSGSERESTLHIRTADSLSEV 429
Qy      426 ASGRQEYSTATDEQ-----AVSTEEYAEEMYD-----GVDDRAGEITAAL 464
Db      430 SNSIDEORTTSEQATTQSTATSVRAGLSDTTALASDAESPVICORESAEELIASL 489
Qy      465 DDIADATDOQVRT-----VBFRVETVG 486
Db      490 EQFQNTAVEQLQSRVASEFTATEDSETAG 518

```

Search completed: August 10, 2004, 15:26:49
 Job time : 16.3923 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

OM protein - protein search, using SW model

Run on: August 10, 2004, 15:22:05 ; Search time 58.609 Seconds
 Perfect score: 2394.
 Sequence: 1 MSNDNTLVTDYRGIDGH.....ATDQQVRVTEEVRETVGKLS 489
 (without alignments)
 2632.505 Million cell updates/sec

Title: US-09-455-978B-2

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Database : SPTREMBL 25:
 1: sp_archea:
 2: sp_bacteria:
 3: sp_fungi:
 4: sp_human:
 5: sp_invertebrate:
 6: sp_mammal:
 7: sp_mhc:
 8: sp_organelle:
 9: sp_phage:
 10: sp_plant:
 11: sp_rabbit:
 12: sp_virus:
 13: sp_vertebrate:
 14: sp_unclassified:
 15: sp_rvirus:
 16: sp_bacteriap:
 17: sp_archaea:

ALIGNMENTS

RESULT 1

ID	P71412	PRELIMINARY;	PRT;	482 AA.
AC	P71412;			
DT	01-FEB-1997 (TREMBLrel. 02, Created)			
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Transducer HtA protein.			
GN				
OS	Halobacterium salinarum.			
OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;			
OC	Halobacteriaceae; Halobacteria; Halobacterium.			
OX	NCBI_TAXID:2242;			
RN	[1] AAB17880.1;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=FLX15;			
RX	MEDLINE=96709786; PubMed=8643458;			
RA	Zhang W., Brooun A., McCandless J., Banda P., Alam M.;			
RT	"Signal transduction in the archaeon Halobacterium salinarium is processed through three subfamilies of 13 soluble and membrane-bound transducer proteins.".			
RT	Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654 (1996).			
RL	EMBL: U75435; AACB17880.1; -;			
DR	PIR: T44973; T44973.			
DR	HSSP: P02942; 1007.			
DR	GO: GO:0016020; C:membrane; IEA.			
DR	GO: GO:0004871; F:signal transducer activity; IEA.			
DR	GO: GO:0005935; P:chemotaxis; IEA.			
DR	GO: GO:0007165; P:signal transduction; IEA.			
DR	InterPro: IPR04089; Chm_axis_transd.			
DR	InterPro: IPR03660; HAMP.			
DR	InterPro: IPR04090; Me_chemotaxis.			
DR	InterPro: IPR000014; PAS_domain.			
DR	PRINTS: PR00260; CHMTRNSDUCR.			
DR	SMART; SM03034; HAMP; 1.			
DR	SMART; SM0283; MA; 1.			
DR	SMART; SM00091; PAS; 1.			
DR	PROSITE; PS50111; CHEMOAXIS_TRANSIDC_2; 1.			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
--	--	--	--	P71412 halobacteri
1	609.5	25.5	482 1 P71412	Qhpj10 halobacteri
2	597	24.9	633 17 Qhpj10	Qphj10 halobacteri
3	597	24.9	805 1 Qpbj10	Qpbj10 halobacteri
4	580	24.2	790 17 Qhr88	Qhr88 halobacteri
5	560.5	23.4	452 1 093775	093775 halobacteri
6	540.5	22.6	777 1 P71415	P71415 halobacteri
7	53.9	22.5	792 1 P71414	P71414 halobacteri
8	53.5	22.4	636 17 Qhqx8	Qhqx8 halobacteri
9	52.9	22.1	804 1 P71416	P71416 halobacteri
10	52.4	21.9	451 1 006022	006022 halobacteri
11	52.3	21.9	544 1 P71409	P71409 halobacteri
12	52.3	21.9	643 17 Qhpq5	Qhpq5 halobacteri
13	521.5	21.8	628 17 Qhrn6	Qhrn6 halobacteri
14	509.5	21.3	789 17 Qhr1	Qhr1 halobacteri
15	507.5	21.2	788 1 091643	091643 halobacteri
16	506	21.1	420 17 Qhpw6	Qhpw6 halobacteri

DR	PROSITE; PS50112; PAS; 1.	DR	GO:0016020; C:membrane; IEA.
SQ	SEQUENCE 482 AA; 50872 MW; ECBEFF9B3374CEC7 CRC64;	DR	GO:0004811; F:signal transducer activity; IEA.
Query Match	25.5%: Score 609.5; DB 1; Length 482;	DR	GO:0006535; P:chemotaxis; IEA.
Best Local Similarity	37.2%; Pred. No. 1.2e-18;	DR	GO:00007165; P:signal transduction; IEA.
Matches	181; Conservative 70; Mismatches 197; Indels 39; Gaps 12;	DR	InterPro; IPR004889; Chmotaxis_transd.
QY	15 NGIDGHALADRIGDEAELAWRLSFTGDDDTMAALAAEQPLFEATADAL-YTDFYDLE 73	DR	InterPro; IPR003660; HAMP.
Db	1.9 NTLDVPAFA--VDAVGAVAW-----DDQIAAL----LETAPEAIGVTDIGEHLN 63	DR	InterPro; IPR004090; Me_chemotaxis.
QY	74 SYERTQDLFANSKTKTVEQLKETQAFLYLGLGRGYDTEYQAQRARIKGTHDVGLGPVVY 133	DR	Pfam; PF0015; MCPsignal; 1.
Db	64 D---DGSRALANKVADTTIDAHEY---DGVGLADESYALLTGDDYYEDTTVAGNTLW 116	DR	PRINTS; PRO0060; CHEMTRNSDUCR.
QY	134 LGAYTRYTGTLLDALADYDVADRGEEAAAVADLDELVARLPLMKLIL----TBDQI--- 184	DR	SMART; SM00304; HAMP; 2.
Db	117 FIATPVHTGEFRGV1-EIVQDR-SSSARYQSEQLFQELVDTDAYDAGRDATVDIA 174	DR	PROSITE; PSS0283; MA; 1.
QY	185 AMDTYIDSYAQRILHDEIDSROELANAVATHVEAPLSSLEATSQDAERTDMARTDDQV 244	DR	PROSITE; PSS0885; HAMP; 2.
Db	175 AEDTLLDEYIQLCRNLTEFGDTLAAHTEVHANDVERLAAQSQVSESSAEIDLSTSQS 234	DR	Complete genome; NW; 64423 AA; C4D802B66EDOFF CRCE4;
QY	245 DRMADVSRREISSYASASVEVASTADDYRTSEDAEAALAQGEAAAADDALATMDTIDEATD 304	DR	Query Match 24.9%; Score 597; DB 17; Length 633;
Db	235 TNVSTVATEVETLSATVQEIASTADEVYDTSATAERLADDGSAAASDADMADVATAAD 294	DR	Best Local Similarity 33.8%; Pred. No. 5.7e-18;
QY	305 GVTAGVEQOLGERAADVESVTGVDDIAETTNMLALNASEAARAGEEGGFAVVADETKA 364	DR	Matches 150; Conservative 76; Mismatches 158; Indels 60; Gaps 7;
Db	295 SVTSDEVALQNRLEDIEVVDVITGIAEQETNMLALNASEAAPGGESEGFAVVADETKA 354	DR	Query Match 24.9%; Score 597; DB 17; Length 633;
QY	365 LABESREQSTRVERLVEQVQMAETEVQDOLDEVNQRIGEVVERVEEAMNETLQB1TDAVE- 423	DR	Best Local Similarity 33.8%; Pred. No. 5.7e-18;
Db	355 LAEPAQSNAGHISLVSIEIQRDATDVTDLVTTDRIPAVQEDANASFEEVTAPEA 414	DR	Matches 150; Conservative 76; Mismatches 158; Indels 60; Gaps 7;
QY	424 - -DAASGMQEVSTATDEQAVSTTEVAEMYGVDDRAIGEAAALDDIADATDQOVRTVEV 481	DR	Query Match 24.9%; Score 597; DB 17; Length 633;
Db	415 TAERASNN--VS DATNEQASAEIIAAMYDETADLADITTAVIDAVSOTEAASAMLIDL 472	DR	Best Local Similarity 33.8%; Pred. No. 5.7e-18;
QY	482 RETVGKL 488	DR	Matches 150; Conservative 76; Mismatches 158; Indels 60; Gaps 7;
Db	473 DESVSEL 479	DR	Query Match 24.9%; Score 597; DB 17; Length 633;
RESULT 2			
Q9HP10	PRELIMINARY; PRT; 633 AA.	DR	Q99J0; PRELIMINARY; PRT; 805 AA.
ID	Q9HP10	DR	Q99J0; PRELIMINARY; PRT; 805 AA.
AC	Q9HP10;	DB	Q99J0; PRELIMINARY; PRT; 805 AA.
DT	01-MAR-2001 (TREMBLrel. 16, Created)	QY	517 GDRVADGETTEAALAAULDDIGQVEAANGSVSISIDATEQAASTEEVTMIDEVTLS 576
DT	(TREMBLrel. 16, Last sequence update)	DR	517 GDRVADGETTEAALAAULDDIGQVEAANGSVSISIDATEQAASTEEVTMIDEVTLS 576
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	QY	458 GETAAALDDIADTQQVRTVEV 481
DE	Htr-3 transducer	DR	458 GETAAALDDIADTQQVRTVEV 481
GN	Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).	DR	577 DRATESQQVSAAEEQAASVSEV 600
CC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;	RESULT 3	
NCBI TaxID	64091;	Q99J0	
OX	[1]	Q99J0	
RN	SEQUENCE FROM N.A.	AC	Q99J0;
RX	SEQUENCE FROM N.A.	DT	01-OCT-2000 (TREMBLrel. 15, Created)
MEDLINE-2050483; PubMed=11016950;	Created)	DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)
NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,	(TREMBLrel. 16, Last sequence update)	DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)
Shukla H.D., Lasky S.R., Baliga N.S., Thorogood V., Sbrosga J.,		DB	Chmotaxis transducer Protein Bast.
Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,		GN	BAST.
Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,		OS	Halobacterium salinarum.
Maddocks D.G., Jablonski P.E., Krebs M.P., Angewine C.M., Dale H.,		OC	Archaea; Euryarchaeota; Halobacteria; Halobacterium.
Iserbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,		OC	Halobacteriaceae; Halobacterium.
Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,		NCBI TaxID=2242;	
Elshabani H., Low T.M., Liang P., Riley M., Hood L., Dassarma S.,		RN	[1]
RT	"Genome sequence of Halobacterium species NRC-1."	RP	SEQUENCE FROM N.A.
PROC	Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).	RC	STRAIN=59;
EMBL	AB005085; AUG200060.1 -.	RX	MEDLINE-20138366; PubMed=10672186;
PIR	HB-336; HB4336.	RA	Kofoeva M.V.; Oesterheld D.;
HSSP	P02942; 1QUT.	RT	"Bast", a membrane-bound transducer protein for amino acid detection in

RT	Halobacterium salinarum.";	RX	MEDLINE=20504483; PubMed=11016950;
DR	NL. Microboli. 35:647-656 (2000).	RA	Ng W.Y., Kennedy S.P., Malairas G.G., Berquist B., Pan M., Sbragna J.,
EMBL	AJ24950; CAB82372.1, -.	RA	Snuika H.D., Lasky S.R., Ballig N.S., Thorsson V., Sbragna J.,
PIR	T48840; T48840.	RA	Swartzell S., Weir D., Hall J., Dahl T.A., Weitz R., Goo Y.A.,
DR	HSSP: P02942; 1Q07.	RA	Leithauer B., Keller K., Cruz R., Danzon M.J., Hough D.W.
GO	GO:0016020; C:membrane; IEA.	RA	Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
GO	GO:004871; P:signal transducer activity; IEA.	RA	Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
DR	DR: GO:006335; P:signal transduction; IEA.	RA	Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
DR	DR: GO:0007165; P:signal transduction; IEA.	RA	Ebbhardt H., Low T.M., Liang P., Riley M., Hood L., DasSarma S.;
InterPro	IPR00409; Chmataxis_transd.	RA	"Genome sequence of Halobacterium species NRC-1."
InterPro	IPR004090; Me_HAMP.	RL	Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
InterPro	IPR004090; Me_chemotaxis.	DR	EMBL: AE005023; AAC19270.1, -.
pfam	PF00672; HAMP; 1.	DR	PIR: BB4238; BB4238.
DR	Pfam; PF00672; MCPsignal; 1.	DR	HSSP: P02942; 1Q07.
PRINTS	PRO0260; CHEMTRNSDUCR.	DR	GO; GO:0016020; C:membrane; IEA.
SMART	SM00304; HAMP; 2.	DR	GO; GO:0004871; P:signal transducer activity; IEA.
SMART	SM00283; MA_1.	DR	GO; GO:0006335; P:chemotaxis; IEA.
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSduc_2; 1.	DR	InterPro; IPR004089; Chmataxis_transd.
DR	PROSITE; PS50885; HAMP; 2.	DR	InterPro; IPR004060; HAMP.
DR	PROSITE; PS50885; HAMP; 2.	DR	InterPro; IPR004090; Me_chemotaxis.
SEQUENCE	805 AA; 84831 MW; 089CA734D3F9DE3D CRC64;	DR	InterPro; IPR004090; Me_Chemotaxis.
Query Match	24.9%; Score 597; DB 1; Length 805;	DR	PFam; PF00672; HAMP; 1.
Best Local Similarity	33.8%; Pred. No. 7..6e-18;	DR	PRINTS; PRO0260; CHEMTRNSDUCR.
Matches	150; Conservative 76; Mismatches 158; Indels 60; Gaps 7;	DR	SMART; SM00304; HAMP; 2.
Qy	42 IDDDTMALAAEQPLFEATADALVTFYDHLEYERT---ODLFANSTKIVQEQLKETQA 97	DR	PROSITE; PS50283; MA_1.
Db	385 LDEDLPGAGFAGSLQSMHTRLEALITLDLEARDAEQTTRKDAEARASERLNRLERRA 444	DR	PROSITE; PS50885; HAMP; 2.
Qy	98 EYLLGLGRGEYDTEYAAQRARIGKIHDLVGLGPDVYLGAYTRYTGYLGLADDVYADRG 157	KW	Complete proteome.
Db	445 -----EYSDEMALAA-----GDLTR-----RLDEDVDSEPM 471	SEQUENCE	790 AA; 84320 MW; 6222D876B5A35BB95 CRC64;
Query Match	24.2%; Score 580; DB 17;	Query Match	24.2%; Score 580; DB 17;
Best Local Similarity	31.4%; Pred. No. 3..9e-19;	Best Local Similarity	31.4%; Pred. No. 3..9e-19;
Matches	160; Conservative 84; Mismatches 190; Indels 76; Gaps 11;	Matches	160; Conservative 84; Mismatches 190; Indels 76; Gaps 11;
Qy	1 MSNDNDNTLVATDVRNGIDGHALADRGILDEAEIARLISFTGIDD--DTMAALA--AEQP 55	Qy	1 MSNDNDNTLVATDVRNGIDGHALADRGILDEAEIARLISFTGIDD--DTMAALA--AEQP 55
Db	286 MSDRAEAIAAGDIDTAATEETRIDEVG-----DURRSFRD1QYE1-LTVAGQDADLAEQD 339	Db	286 LPEATA-----DALVTDFYD----HLESYERTQDLPNSTKIVQEQLKETQAEVLLG 102
Qy	56 LPEATA-----DALVTDFYD----HLESYERTQDLPNSTKIVQEQLKETQAEVLLG 102	Qy	56 LPEATA-----DALVTDFYD----HLESYERTQDLPNSTKIVQEQLKETQAEVLLG 102
Db	340 -FDADALDKSVPGRIGESLEIMHWDLETAIALEOQETAOSRKAEQSRE-EARALA- 396	Db	340 -FDADALDKSVPGRIGESLEIMHWDLETAIALEOQETAOSRKAEQSRE-EARALA- 396
Qy	103 LGRGEYDTEYAAQRARIGKIHDLVGLGPDVYLGAYTRYTGYLGLADDVYADRGEEAAA 162	Qy	103 LGRGEYDTEYAAQRARIGKIHDLVGLGPDVYLGAYTRYTGYLGLADDVYADRGEEAAA 162
Db	397 -----AALBSAQDIRKETVEAAD-----GDTQRLETDTDEBSMA 432	Db	397 -----AALBSAQDIRKETVEAAD-----GDTQRLETDTDEBSMA 432
Qy	163 AVDELVARYRFPLMLKLJLTFDQIAMIPTYDSYAQRLHDEDSRQELANAVA -- THVEAPL 219	Qy	163 AVDELVARYRFPLMLKLJLTFDQIAMIPTYDSYAQRLHDEDSRQELANAVA -- THVEAPL 219
Db	433 AI-----ATALNSLLELELTGTHRIQRFSDVASSDHTTSA 470	Db	433 AI-----ATALNSLLELELTGTHRIQRFSDVASSDHTTSA 470
Qy	531 SVGSGRSQASDAIEMNADEVQATEETVQDLMQAEQDDEMTIEGEIVTLDDAEQTSML 590	Qy	531 SVGSGRSQASDAIEMNADEVQATEETVQDLMQAEQDDEMTIEGEIVTLDDAEQTSML 590
Db	340 NASTEAAARGAARGAGEGFAYVADBEVKALAAESREQSTREELVEOMQAETEETVQDLMQEVNQ 399	Qy	340 NASTEAAARGAARGAGEGFAYVADBEVKALAAESREQSTREELVEOMQAETEETVQDLMQEVNQ 399
Qy	591 NASTEAAARGAARGAGEGFAYVADBEVKALAAESREQSTREELVEOMQAETEETVQDLMQEMGD 650	Db	591 NASTEAAARGAARGAGEGFAYVADBEVKALAAESREQSTREELVEOMQAETEETVQDLMQEMGD 650
Db	01-MAR-2001 (TREMBLrel. 16, Created)	Qy	400 RIGEGVEREAMETIQETDAEDASGMQEVSSTATDEQAVSTEEVAEMVTDGVDDRAGE 459
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)	Db	651 RLSEGKSTVTDVETDIVERANGQVOTINTATDEQATTTEEVVMDEVSSISDD 710
DB	HtrB transducer.	Qy	460 IAAALDDIADATDQCVRTVEBV 481
GN	HtrB OR VNG0812G.	Db	711 TTARAEDAAAEQTAGTASLTEVTNRQDL 740
OS	Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).	SEQUENCE FROM N.A.	SEQUENCE FROM N.A.
OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;		
NCBI_TaxID	64091.		
RN	[1]		
RP			

Db	612 RADQDGFFAVVADEVKOLADESKALAAEMQALVAEVRAQETSVAAAMDRIQERVSDGV	671	Db	556 AATESERGQELGAVAVERLERIBATADSAYERVTALEBAVDAGHTVGVITDIAEQTNML	615
Qy	407 RVEEAMETILOEITDAVEDAASGMQEVSSTATDEQAVSTTEVAEMYDGVDDRASHIAANALD	466	Qy	338 ALNASIEEARAGAGEGAVVABVRALAEESQEOSTREVELYEQMQUEETEYDQLDEV	397
Db	672 TVSETERSSISDIAGRIAEBDTGQBEISNAMMDQASVSDVTAVGDYRALGEETATEAES	731	Db	616 ALNNIEEARADSGDGPAAVDAEVDKLADEVKESEATELETLYDVQADTVAADMSEL	675
Qy	467 IADATDQYRTVEEV 481		Qy	398 NORIGEGYVERVERAMETILOEITDAVEDAASGMQEVSSTATDEQAVSTTEVAEMYDGVDDRA	457
Db	732 TRDAAAEQTTLSDV 746		Db	676 GDRDVAGSETTEALAAVDIGQVEAANGSVSISDATADEQASTEEVVMTIDEVTDL	735
RESULT 7			Qy	458 GETAAALDDIADATDQYRTVEEV 481	
P71414	ID P71414 ; PRELIMINARY ;	PRT ; 792 AA.	Db	736 DRITATCQSVAEAEQASVSEV 759	
AC	P71414 ;				
DT	01-FEB-1997 (TREMBLrel. 02, Created)		Q9H0X8	PRELIMINARY ;	
DT	01-FEB-1997 (TREMBLrel. 02, Last sequence update)		ID Q9H0X8 ;	PRT ;	636 AA.
DE	Transducer HtC protein.		AC Q9H0X8 ;		
GN	HtC.		DT 01-MAR-2001 (TREMBLrel. 16, Created)		
OS	Halobacterium salinarum.		DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
OC	Euryarchaeota; Halobacteria; Halobacteriales;		DE Htr15 transducer.		
OC	Halobacteriaceae; Halobacterium.		GN HTR15 OR VNGC09586.		
OX	NCBI_TAXID=2242;		OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).		
RN	[1] -		OC Archaea: Euryarchaeota; Halobacteria; Halobacteriales;		
RP	SEQUENCE FROM N.A.		OC Halobacteriaceae; Halobacterium.		
RC	STRAIN=Flx15;		OX NCBI_TAXID=4091;		
RX	RESEQUENCE FROM N.A.		RN [1] -		
RA	Zhang W., Brown A., McCandless J., Banda P., Alam M.;		RA RA		
RT	"Signal transduction in the archaeon Halobacterium salinarium is		Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,		
RT	processed through three subfamilies of 13 soluble and membrane-bound		Shukla H.D., Iasky S.R., Morison V., Sbrigno J.,		
RL	transducer proteins." / Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654(1996).		Swartzell S., Weir D., Hall J., Dahl T.A., Welti Y.A.,		
DR	HSSP; P02942; AAB17882.1; -.		RA Leithäuser B., Keller K., Cruz R., Danson M.J., Hough D.W.,		
DR	RT "Signal transduction in the archaeon Halobacterium salinarium is		Maddock D.G., Tablonski P.E., Krebs M.P., Argovine C.M., Dale H.,		
GO; GO:0016120; C:membrane; IEA.	processed through three subfamilies of 13 soluble and membrane-bound		Isenburger T.A., Peck R.F., Spudich J.L., Jung K.,		
DR	RT "Signal transduction in the archaeon Halobacterium salinarium is		Alam M., Freitas T., Hou S., Dennis P.P., Omer A.D.,		
GO; GO:004871; F:signal transducer activity; IEA.	processed through three subfamilies of 13 soluble and membrane-bound		Ebbhardt H., Lowe T.M., Liang P., Riley M., DasSarma S.,		
DR	RT "Signal transduction in the archaeon Halobacterium salinarium is		RT Genome sequence of Halobacterium species NRC-1."		
GO; GO:0093935; P:chemotaxis; IEA.	processed through three subfamilies of 13 soluble and membrane-bound		Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).		
DR	RT "Signal transduction in the archaeon Halobacterium salinarium is		DR HPI; AE005032; AGI91381.1; -.		
GO; GO:0001765; P:signal transduction; IEA.	processed through three subfamilies of 13 soluble and membrane-bound		DR PIR; A84252; AB4252.		
DR	RT "Signal transduction in the archaeon Halobacterium salinarium is		DR HSSP; P02942; 1QUT.		
InterPro; IPR004089; Chmtaxis_transd.	processed through three subfamilies of 13 soluble and membrane-bound		DR GO; GO:00016020; C:membrane; IEA.		
InterPro; IPR003660; HAMP.	processed through three subfamilies of 13 soluble and membrane-bound		DR GO; GO:0000155; P:two-component sensor molecule activity; IEA.		
InterPro; IPR004090; Me_chmotaxis.	processed through three subfamilies of 13 soluble and membrane-bound		DR GO; GO:0006935; P:chemotaxis; IEA.		
InterPro; IPR006722; HAMP; 1.	processed through three subfamilies of 13 soluble and membrane-bound		DR InterPro; IPR004089; Chmtaxis_transd.		
Pfam; PF00015; MCPsignal; 1.	processed through three subfamilies of 13 soluble and membrane-bound		DR InterPro; IPR004090; Me_chmotaxis.		
PRINTS; PRO0260; CHEMTRNSDUCR.	processed through three subfamilies of 13 soluble and membrane-bound		DR InterPro; IPR00700; PASassoso.C.		
DR SMART; SM00304; HAMP; 2.	processed through three subfamilies of 13 soluble and membrane-bound		DR SMART; SM00091; PAS; 1.		
DR SMART; SM00283; MA; 1.	processed through three subfamilies of 13 soluble and membrane-bound		DR TIGR_PAMS; TIGR0229; sensory_box; 1.		
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDC_2 ; 1.	processed through three subfamilies of 13 soluble and membrane-bound		DR PROSITE; PS50111; CHEMOTAXIS_TRANSDC_2 ; 1.		
DR PROSITE; PS50885; HAMP; 2.	processed through three subfamilies of 13 soluble and membrane-bound		DR PROSITE; PS50113; PAC; 1.		
SQ SEQUENCE 792 AA; MW; 745D3693F3EEF9F9 CRC64;	processed through three subfamilies of 13 soluble and membrane-bound		DR PROSITE; PS50112; PAS; 1.		
Query Match 22.5%; Score 539 ; DB 1; Length 792;	processed through three subfamilies of 13 soluble and membrane-bound		KW Complete proteome.		
Best Local Similarity 32.2%; Pred No. 2.e-15; Mismatches 76; Indels 62; Gaps 7;	processed through three subfamilies of 13 soluble and membrane-bound		SQ 636 AA; 67349 MW; 98D2E415295DAD14 CRC64;		
Matches 143; Conservative 76; Mismatches 163; Indels 62; Gaps 7;	processed through three subfamilies of 13 soluble and membrane-bound		Query Match 22.4%; Score 535 ; DB 17; Length 636;		
DR 374 LDEDLPGAFFGASLSQNH-RFLALITDLEAERDAEQTKRQDEEARASEINERJERRAA	processed through three subfamilies of 13 soluble and membrane-bound		Best Local Similarity 29.7%; Pred No. 2.4e-15;		
DR 460 QDIAEAENDMGDV-----EATLAQVRSIAAVADSDTDAEAE-----	processed through three subfamilies of 13 soluble and membrane-bound		Matches 149; Conservative 99; Mismatches 191; Indels 63; Gaps 11;		
Qy 98 BYLLGLGRGYDTEYYAQRARIGKTHDVLGIPDQYVLAGTRYTGTLADDVYADRG 157	processed through three subfamilies of 13 soluble and membrane-bound		Qy 17 IDGHALDRIGLDNEAIANLISFCLDDTMALAAEQLFPEAT---ADALVTD---		
Db 433 -----EYSDDEMAAAA-----GDLTR-----FLDEDVDSGM 459	processed through three subfamilies of 13 soluble and membrane-bound		Db 165 VDDDAILDSTGMP---VFMPLDATADEQASVYAWNETIEELTGCSKADAMGMDHASEA 214		
Qy 158 BEAAAAYDLYARFLPMLKLTFDQQTAMIDYDSTAQRHLDEIDSRLQELANAVATHVEA	processed through three subfamilies of 13 soluble and membrane-bound		Qy 278 AEALAQGEAADDALATMMTIDEATDGVTAGVEQLGERAADVESVTGVDDIAEQTNML 337		
Db 460 QDIAEAENDMGDV-----EATLAQVRSIAAVADSDTDAEAE-----	processed through three subfamilies of 13 soluble and membrane-bound		Qy 68 FYDHLSEYERTQDLFANSRKTRVEQLKETOREYLGLGRGYDTEYAQRARIGKIHDLG 127		

Db	215 FYPDDRRVKTAD-----KVISSPRSAAA-----EFIERSPEKAQIYADTSVM-	258	Qy	89 VEQLKETQAETYLGLGRGEYDTEYAAQR-----ARIGKTHDVGLGLPVDYLGAYTRY 140	
Qy	128 LGPDVVILGAYTRYTGLIDALLADD-----VVADEGEAA-AAADEVLFPLML-	175	Db	351 VQSLSAARE-----IEAGNYDWDVATSRREDETGQLASIGSMRDLVTOAE---AAREQA 403	
Db	259 ---TDQGTDRHIRKASPPIPPDDGELLAYAETIEDRTEDVRADEVELYDELSTTID	314	Qy	141 YTGLLDALAD----DLYVADRGEEAAADEVLFPLMLKLTFDQOIAMDTYIIDSQAQR 196	
Qy	176 -----KLIFDQOIAMDTYIDSQAQRHEDBDSROELANAVATEAPLSSLEATSQ	227	Db	404 TEAQDAEADARRAEDERADEAALAE-----LEADER 444	
Db	315 ALSSGQLSKRASPEHEGGINQEVISVVALNAGDQFPLRVLFCVQDGQTQELADTIEPAT	374	Qy	197 LHDEI-----DSR-----QELANAVAT--HV 215	
Qy	228 DVAERTDTMRARTDDQVDRMADVSEISSVSASVEVEYASTADDVTRTSEDAFAALAQOGEA	287	Db	445 YSDVIVGGVADGDILTRMPADDTNEMAAIAASFNSQWEHTIDIOQFADAVATASEA 504	
Db	375 DATDIADTV----SSONEMLSSAANEMENFSASMQEYAAASSQVASAEEQADAESGLE	430	Qy	216 EAPLSSLEATSDQVAERTDTMARTDDQVDRMADVSEISSVSASVEVASTADDVTRT	275
Qy	288 AADDALATMTDIDBEATDQVTAGVQQLGRRAADESITVGVIDDIAEQTMNLAINALEAR	347	Db	505 EVGAADAERASQSVESQVEIAGADEQRMNLTDSSEMIDISAEEVAS-ADSYAEHS 563	
Db	431 ASEGANQATNEVIDISDLMESVKLESRMDEEDVVEVIAVADQTMNLAINALEAR	490	Qy	276 EDARALACQGEAAADDALATMTDIDEATDGTAGVQOLGERAADVESVTGVIDDIAEQTN 335	
Qy	348 AGEAGEGFAVVADEVKALAEESREQUESTEELVEQMQAEETEVTQDOLDEVNRIGEGER	407	Db	564 HQTAEIARDGEQTAEDAATESLTQEALDATVNVEALDDQMAEISEIVDLSDIAEQTN 623	
Db	491 AGFAGSGFAVVADEVKEANE TREHTEIAGSISDVOQANEITVLAVESSHQIHRAGDE	550	Qy	336 MLLANLNASTEAAAGAESEGFQAVADEYKALAESEREOSTRVELVEQMQAEETEVTQDQ 395	
Qy	408 VEEFAMETLQETIDAVEDAASGMQEVSSTATDEQAVSTPFRVAEMDVGDDRAGEAAALDDI	467	Db	624 MLLANLNASTEAAAGKSGDFQAVADEYKDFASETQEADIRRITERQSQTTAVEAR 683	
Db	551 IDDLATAEETATSVDEEATGTTTEVARANDEQATSTEVLYVITLEDVQOCQEEAAASDRI	610	Qy	396 EVNRIGEGFQVEEAMETLQETIDAVEDAASGMQEVSSTATDEQAVSTEEVAMVGDYDD 455	
Qy	468 ADATDQQRTVVEEVRTVGKLS	489	Db	684 AAEESMSMAGIDAVEEVDAFTAVSDHSDETDIGVQEISDRTDDQAASTEEAVSMTEEAVAD 743	
Db	611 VSATQEQTAVSQLSERVDKLT	632	Qy	456 RAGETIAAALDDIADATDQVRITVEEVRTVGKLS 489	
			Db	744 LSDSTAGEAQVSATAAEQQASMSIEISDTVESTIS 777	

RESULT 9

P71416	PRELIMINARY;	PRT;	804 AA.	RESULT 10
AC	P71416;			006022 PRELIMINARY; PRT; 451 AA.
DT	01-FEB-1997 (TREMBLrel. 02, Created)	ID	006022	PRELIMINARY;
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)	AC	006022;	
DE	Transducer HtF protein.	DT	01-JUL-1997 (TREMBLrel. 04, Created)	
GN	HtF.	DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
OS	Halobacterium salinarum.	DE	Soluble transducer protein HtH.	
OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;	OS	Halobacterium salinarum.	
CX	Halobacteriaceae; Halobacteria; Halobacterium.	OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;	
NCPBI_TaxID=2242;		OX	Halobacteriaceae; Halobacteria; Halobacterium.	
RN	[1]	NCBI_TaxID=2242;		
RP	SEQUENCE FROM N.A.			
RC	SEQUENCE FROM N.A.			
RX	STRAIN=Fix15;			
RX	MEDLINE=96209786; PubMed=8643458;			
RX	Zhang W., Brooun A., McCandless J., Banda P., Alam M.;			
RT	"Signal transduction in the archaeon Halobacterium salinarum is processed through three subfamilies of 13 soluble and membrane-bound transducer proteins."			
RT	Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654 (1996).			
RL	DR EMBL; U75439; AAA17884.1; -.			
DR	PIR; T44606; T44606.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0004871; P:signal transducer activity; IEA.			
DR	Zhang W., Brooun A., McCandless J., Banda P., Alam M.;			
DR	"Signal transduction in the archaeon Halobacterium salinarum is processed through three subfamilies of 13 soluble and membrane-bound transducer proteins."			
DR	Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654 (1996).			
DR	InterPro; IPR003660; HAMP.			
DR	InterPro; IPR00672; HAMP; 1.			
DR	Pfam; PF00015; MCPSignal; 1.			
DR	SMART; SM00304; HAMP; 2.			
DR	SMART; SM00283; MA; 1.			
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSduc_2; 1.			
DR	PROSITE; PS50085; HAMP; 1.			
SEQUENCE	804 AA;	0D57284683BD18B CRC64;		
SQ	Score 529, DB 1; Length 804;			
Query Match	22.1%			
Best Local Similarity	30.6%			
Matches	Conservative 76;	Pred. No. 5.9e-15;		
Matches	Mismatches 159;	Indels 80;	Gaps 9;	
Query Match	21.9%;	Score 524;	DB 1;	Length 451;

Best Local Similarity Matches	30.9%; 158; Conservatve	Pred. No.	4.9e-15;	DR PRINTS; PRO0260; CHEMTRNSDUCR.
Qy	10 TADVRNGIDGHALADRI-----GLDEAIA-----WRLSPTGIDDDTMALAA	Indels	98;	Gaps 16;
Ds	5 SSMGGGEATGEHLADELCEAYLGDNEDGGDELQRLSRERDWKHMN-----QLVA		56	DR SMART; SM00304; HAMP; 1.
Qy	53 EQP--LFEPATADALVTD----FYDHESY----ERTQDLFANSTKTEVOLKETQAEYL 100			DR SMART; SM00283; MA; 1.
Ds	57 EPEGILLITAAQRARIGKTHDVLGLGPDVYLGYATRYYTGLLDALADYYVADRGEEA 160			DR PROSITE; PS50111; CHEMOTAXIS_TRANSUC_2; 1.
Qy	101 LGHGRGEYDTEAAQRARIGKTHDVLGLGPDVYLGYATRYYTGLLDALADYYVADRGEEA 160			DR PROSITE; PS50885; HAMP; 1.
Ds	112 VRTGDTVEEE-----PHDVTPLSCLCYHGVPRLAPFG-----DVVGSGF--V 152			SEQUENCE 544 AA; 56922 MW; 6630D80E43AFC471 CRC64;
Qy	161 AAAVDELVAFRUPMLKLITFDQIAMDUTYIDSYAQRILHDEISRSQELANAVATHVEAPLS 220			Query Match 21.9%; Score 523.5.; DB 1; Length 544;
Ds	153 VPDISEKVN-----QRELHD-----LHEVTSINVGEHLS 182			Best Local Similarity 29.4%; Pred. No. 6.4e-15; Mismatches 93; Conservative 93; Indels 51; Gaps 10;
Qy	221 SLEATSQDV--AERTDTMARTDDQVDRMDSRETTSSASVSEEVASTADDVRRTSED 277			Matches 141; Conservative 141; Indels 51; Gaps 10;
Ds	183 ELSESTIKVGPSARETPEAGK--EERMEGFADKVSQSNTTIEAASAEVSQASQR 239			28 LDDEAETIARWLRSFTGIDDDTMALAAEOPLEFEATDALVTF---VDHLESVERTQDPLFA 83
Qy	278 ABLAACQGEAAADDALATMDIDEATDGTVATGEQVOLGERAADEVSYTGIVTDIAEQTNML 337			64 LEKAGVATQI-FNLATGPLMATTVAFAIGGTTALIVEDMEALVEERAQPRQEABERA 122
Ds	240 AQRATEGEQTAAETIDRMGAQESAERVNDTIDGLTSQADEMSETIDAINDIACINML 299			84 NSTKTVFQLKETQAEYVLLGLGRGEYDT-EYAAQRARIGKTHDVLGLGPDVYLGYATRYYT 142
Qy	338 ALNASTIAARAAGEAGEFAVVADEVKALESEREQSTRVELVEQMQAEETVQDOLDEV 397			123 EAERARAKEAQXQAE-----AERQAEAOQSQTGTGTQREIEQQAAD--LESQTEVG 173
Ds	300 ALNASTIAARAAGEAGEFAVVADEVKALESEREQADEMSETIDAINDIACINML 359			143 GLLDALAD-----DVVADRG-----EAAADEVLFPLMKLJTFDQJIAMDTYDSY 193
Qy	398 NORIGEGYVERSEAMETLQEIITDAVEDAAGMNEQEVSTATDBAVSTEVAlENVDGYDDRA 457			Db 174 ATLEASASDGDILTARVATTDNAEIAAVATVNDM-----LTMTERTI----- 215
Ds	360 TTIEEIAITAATVRETLDSLQLERNADETATGYKEVAGARD-HAASTEQVQARRTDEAVDKL 418			Qy 254 ISSVSASVEEASTDDVVRRTSEDABALAQCQGEAAADDALATMDIDEATDGTVATGEYQL 313
Qy	458 GEJAAALDDIADATDQOVTYBEVRETVGKL 488			Db 271 MDSYSATVEEVATATSVADTAADTDTDVATQKQTEDAIDIAVQETMOTTVANNDAL 330
Ds	419 TELEDRDLNLSQIASEOHDRVAEEDMVDEL 449			Qy 314 GERAAEVESVTGVIDDTAECOTMLNALNASTEAARAGEA---GEGFAVVADEVKALAEES 369
RESULT 11				Db 331 EDLTIEDDDIALISDAEQTMNLANANTAEAAASGGGTNGDGAVVADDEVKEATES 390
P71409	PRELIMINARY;	PRT;	544 AA.	Qy 370 REQSTRVBEELVQMQAEETVTDQLEDEVNRIGEGERVEREAMETIQETDAVEDASGM 429
AC	P71409;			Db 391 QRSKAKIAELIEEVQQQTATVEELRAVEQRVNDGAABEVETVDAFGAVTENIOETTDGV 450
DT	01-FEB-1997 (TREMBLrel. 02, Created)			Qy 430 QEVSTATDEQAVSTEVAEAMVVDGDDRAGETAALLDIADATDQQRVTREEVRETRYGKLS 489
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			Db 451 QEISQAMDEQRQRSERVVSSVDDIATISQATADRAENSASEEQTASITEVTGSISLAA 510
DE	Transducer Ht1 protein.			RESULT 12
GN				Q9HPQ5 PRELIMINARY; PRT; 643 AA.
OS	Halobacterium salinarum.			ID Q9HPQ5 PRELIMINARY; PRT; 643 AA.
OC	Archaea; Buryarchaeota; Halobacteria; Halobacteriales;			AC Q9HPQ5 PRELIMINARY; PRT; 643 AA.
NCBI_TAXID=2242;	NCHB			DT 01-MAR-2001 (TREMBLrel. 16, Created)
RN				DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
RC	SEQUENCE FROM N.A.			DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
RX	STRAIN=F1X15; MEDLINE=96209766; PubMed=8643458;			DB Htr8 transduce.
RA	Zhang W., Brooun A., McCandless J., Banda P., Alan M.;			GN HTR8 OR VNG1523G.
RT	"Signal transduction in the archaeon Halobacterium salinarum is			OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
RT	processed through three subfamilies of 13 soluble and membrane-bound			OC Archaea; Buryarchaeota; Halobacteria; Halobacteriales;
RL	transducer proteins";			OC Halobacteriaceae; Halobacterium;
DR	Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654 (1996).			OC NCBII_TaxID=64091; NCBI_TaxID=64091;
DR	RT "Signal transduction in the archaeon Halobacterium salinarum is			RN [1] -
RT	processed through three subfamilies of 13 soluble and membrane-bound			RP SEQUENCE FROM N.A.
RT	transducer proteins";			RX MEDLINE=20504483; PubMed=11016950;
RT	proc. Natl. Acad. Sci. U.S.A. 93:4649-4654 (1996).			RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., Sbrrogna J., Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Weir D., Dahl T.A., Weilert R., Gooy A., Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W., Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Swarzell S., Weir D., Hall J., Dahl T.A., Weilert R., Gooy A., Isenbarger T.A., Peck R.F., Kohlschroder M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S., RT "Genome sequence of Halobacterium species NBC-1"; PL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
DR	PIR; T44938; T44938;			DR PIR; H84105; H84305.
DR	GO; GO:0016020; C:membrane; IEA.			DR PIR; H84305; H84105.
DR	GO; GO:000871; F:signal transducer activity; IEA.			DR PIR; H84105; H84305.
DR	GO; GO:0006935; P:chemotaxis IEA.			DR PIR; H84305; H84105.
DR	GO; GO:0007165; P:signal transduction; IEA.			DR PIR; H84105; H84305.
DR	InterPro; IPR004089; Chmotaxis_trans.			DR PIR; H84105; H84305.
DR	InterPro; IPR003660; HAMP.			DR PIR; H84105; H84305.
DR	InterPro; IPR004090; Me chemotaxis.			DR PIR; H84105; H84305.
DR	Pfam; PF00015; MCPSignal; 1.			DR PIR; H84105; H84305.

DR	GO:0016020; C:membrane; IEA.	RA	Ng W.Y., Kennedy S.P., Mahairas G.G., Berquist B., Pan M., Shukla H.D., Lasky S.R., Ballig V., Thorson V., Sbranga J., Swartzell S., Weil D., Hall J., Dahl T.A., Walti R., Goo Y.A., Leithäuser B., Keller K., Cruz R., Danzon M.J., Hough D.W., Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H., Isenburger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H., Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D., Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.; RT "Genome sequence of Halobacterium species NRC1." RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR	GO:000935; P:chemotaxis; IEA.	RL	AEG0010; AACG1912-1; -.
DR	GO:0007165; P:signal transduction; IEA.	DR	AE00010; AACG1912-1; -.
DR	InterPro; IPR004089; Chmrtaxis_transd.	DR	F84219; F84219.
DR	InterPro; IPR004060; HAMP.	DR	
DR	InterPro; IPR004090; Me chemotaxis.	DR	
DR	Pfam; PF00015; MCPsigna_1.	DR	
DR	PRINTS; PRO0260; CHEMTRNSDUCR.	DR	
DR	SMART; SM00304; HAMP; 1.	DR	
DR	SMART; SM00283; MA; 1.	DR	
DR	PROSITE; PS50111; CHENOTAXIS_TRANSduc_2; 1.	DR	
DR	PROSITE; PS50885; HAMP; 1.	DR	
KW	Complete proteome.	DR	
SQ	SEQUENCE 643 AA; 67290 MW; 1D53F2B18PA02401 CRC64;	DR	
QY	8 LIVPADVNGIDGHALADPAGLDEAEIINWRLSLFTGIDDDTMALAREQPLFATADALVTD 67	DR	GO:0008935; P:chemotaxis; IEA.
DB	156 LYVVVFETHGFGMINAERVNNTAAINNPWVNGIIGAFVFLLAG-----ALMAN 205	DR	GO:0007165; P:signal transduction; IEA.
QY	68 FYDHLESVERTODLFANSTKTVQLKEKTAQEAYLGLLGRGEYDEYA----- 113	DR	InterPro; IPR004089; Chmrtaxis_transd.
DB	206 WYSTERSEASQKLERBRKRQAKQVEELBAR----QAEITAEKAFAKRIKADAFAARE 259	DR	InterPro; IPR004060; HAMP.
QY	114 --AQARAIKGKHDVLGIGDPDYLGAYTRYGTGLDALAD-DVADRGEEAAAVDELVA 169	DR	InterPro; IPR004090; Me chemotaxis.
DB	260 AAEQQREVAAILNERLEMTANTYGAANARAADDLSVYRLDPVENDAMAAATASFNML- 318	DR	InterPro; IPR004091; CHEMOTAXIS_TRANSduc_2; 1.
QY	170 RELPMLKLTDQIΛNTYIDSQAQNHLDEIDSQRLNEA 222	DR	Complete proteome.
DB	319 -----DETTETIREIQAYASDVAAASEDADAGTEI 349	DR	PF00015; MCPsigna_1.
QY	223 EATSQDVAYERDTMRARTDDQVDRMADSREISSLVSASVEEVASTADVRRPSEDALA 282	DR	PRINTS; PRO0260; CHEMTRNSDUCR.
DB	350 EDASGQSEVTPYQIAAGDQEKEKLYTSGEMTDLSAIEEYASAISVAEFSHEATAVA 409	DR	SMART; SM00283; MA; 1.
QY	283 QOGEAAAADDALATMTDIDEATDGTVAEQLEGRERAADVYESVTCTVIDIAEQTNMLAINA 342	DR	PROSITE; PS50111; CHEMOTAXIS_TRANSduc_2; 1.
DB	410 GDSEQTAQQAIDSRTVQSASESTVQNTYALDDQLAESIEVLDLISVDAEQTNMLAINA 469	DR	PS50885; HAMP; 2.
QY	343 IEARAGFAGEGFAYVADEVKALAAESEQSTREVVEWQMAETETVDOLDEVNRIG 402	DR	Complete proteome.
DB	470 IEARAGFAGEGFAYVADEVKALAAESEQSTREVVEARTADSSVQ 529	DR	SEQUENCE 643 AA; 67290 MW; 1D53F2B18PA02401 CRC64;
QY	403 EGVERVERBAMETIQLITDAVEDAASGMQEVSSTATDEQAVSTEEVAEWGDVDRAGEIAA 462	DR	SDVYAKERDTMDARTDDQVDRMADVSEISSYSAVSPEVASTADPVERTSEDAEALAAQG 285
DB	530 DAISAVDYYDRAFTGVAAEAAQDQVTDQASGIDEALVADIDQMQAVTTEARTADSSVQ 529	DR	SMSVKSTIEQISAGAENONEKLTAAAGEMTDLSATVIAESTNNVQAQEDAAMGDG 408
QY	463 ALDIADIATDQOVRTVEEVRTEVGK 488	DR	286 EAMADALALATMDIDATEATDGTVAEQLEGRERAADVYESVTCTVIDIAEQTNMLAINA 345
DB	590 DAQQASTIAEQQTAAATISENTAA 615	DR	409 QAATDATAEMDAIESATQTVEMQSIQEEVGRISDLVTMDDIASQTDMALAINA 3468
QY		DR	346 ARAGEAGEGFAYVADEVKALAAESEQSTREVVEEVLQEQAEETEVQDLEVNQRIGEGV 405
DB		DR	469 ENASTSGDFGAVVADEVKSLAEDATATADEVESLIAVEASTSALADEMYMRDVGDGQ 528
RESULT 13		QY	406 ERVBEAMTLOETDAEDAAQMCEYSTATDEQAVSTEEVAEMVGDYDDRAEIAAALD 465
ID Q9HRN6	PRELIMINARY;	PRT;	529 DTITDAALEDIVRDVDAQAGQSOSAQL 488
AC Q9HRN6;		ID Q9HR1	PRELIMINARY;
DT 01-MAR-2001 (TREMBLrel. 16, Created)		AC Q9HR1;	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)		DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DB Htr16 transducer.		DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
GN Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).		DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)	
OS Archaea; Euryarchaeota; Halobacteria; Halobacteriales;			
OC Halobacteriaceae;			
OX NCBI_TaxID=64091;			
RN SEQUENCE FROM N_A.			
RELINE=20504483; PubMed=11016950;			
RELINE=20504483; PubMed=11016950;			
RESULT 14			
Q9HR1			
ID Q9HR1			
AC Q9HR1;			
DT 01-MAR-2001 (TREMBLrel. 16, Created)			
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DB Htr16 transducer.			
GN Halobacterium sp. OR VNG0614G.			
OS Archaea; Euryarchaeota; Halobacteria; Halobacteriales;			
OC Halobacteriaceae;			
OX NCBI_TaxID=64091;			
RN SEQUENCE FROM N_A.			
RELINE=20504483; PubMed=11016950;			
RELINE=20504483; PubMed=11016950;			

DE	Htr6 transducer.	Db	702 QADATQSYYRRVDDVADISQHYTEDAEQVSAAEEQSASVAETARSADDL 751
OS	Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).		
OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;		
OX	NCBI_TAXID=64091;		
RN	[1]		
RP	SEQUENCE FROM N.A.	RESULT 15	
RX	MEDLINE=20504483; PubMed=11016950;	O93643	PRELIMINARY;
RA	Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,	ID	O93643;
RA	Shukla H.D., Lasky S.R., Baliga N.S., Shrosgna J.,	AC	O93643;
RA	Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,	DT	01-MAY-1999 (TREMBLrel. 10, Created)
RA	Leithauer B., Keller K., Cruz R., Dawson M.J., Hough D.W.,	DT	01-JUN-2003 (TREMBLrel. 10, Last sequence update)
RA	Maddock D.G., Jablon P.E., Krebs M.P., Angevine C.M.,	DE	Transducer HtrVI.
RA	Isenburger T.A., Peck R.F., Pohlschroder M., Spudich J.L.,	GN	HtrVI.
RA	Alam M., Freitas T., Hou S., Daniels C.J., Dennis F.P., Omer A.D.,	OS	Halobacterium salinarum.
RA	Ebbhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;	OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
RL	"Genome sequence of Halobacterium species NRC-1."	SEQUENCE FROM N.A.	Halobacteriaceae; Halobacterium.
EMBL	AEP05021; AE05021;	[1]	NCBI_TAXID=2242;
PIR	EB4236; EB4236.	RN	RN
DR	PIR; EB4236; EB4236.	RA	Brown A., Villalbaica F., Freitas T., Hou S., Alam M.;
DR	HSSP; P02942; IQU7.	RT	"Structure, Function and Evolution of Transducer Family from the Archaeon Halobacterium salinarium."
DR	GO; GO:0016020; C:membrane; IEA.	RT	Archaeon Halobacterium salinarium."
DR	GO; GO:0004871; C:signal transducer activity; IEA.	RL	J. Bacteriol. 0-0(1997).
DR	GO; GO:0006335; P:chemotaxis; IEA.	DR	EMBL; AF036231; AAD02052.1; -.
DR	GO; GO:0002165; P:signal transduction; IEA.	DR	PIR; TA4262; TA4262.
DR	InterPro; IPR004089; Chmtaxis_trans.	DR	HSSP; P02942; TA4262.
DR	InterPro; IPR001660; HAMP.	DR	GO; GO:0005935; P:chemotaxis; IEA.
DR	InterPro; IPR004090; HAMP.	DR	GO; GO:00007165; P:signal transduction; IEA.
Pfam	PF000672; HAMP; 1.	DR	InterPro; IPR004019; Chmtrans_Transd.
DR	InterPro; IPR004090; HAMP.	DR	InterPro; IPR003660; HAMP.
Pfam	PF00015; MCPsignall; 1.	DR	InterPro; IPR004090; Me_Chemotaxis.
PRINTS	PRO00260; CHEMTRNSDUCR.	DR	Pfam; PF000572; HAMP; 1.
SMART	SM00283; MA; 1.	DR	PRINTS; PRO00260; CHEMTRNSDUCR.
SMART	SM00304; HAMP; 2.	DR	SMART; SM00283; MA; 1.
PROSITE	PS50011; CHEMOTAXIS_TRANSDIC_2.	DR	PROSITE; PS50011; CHEMOTAXIS_TRANSDIC_2.
PROSITE	PS50011; CHEMOTAXIS_TRANSDIC_2.	DR	PROSITE; PS50089; HAMP; 1.
PROSITE	PS50085; HAMP; 2.	DR	PROSITE; PS50089; MW; 83868 MW; COD0839DDDBFCCE CRC64;
KW	Complete Proteome.	SQ	SEQUENCE 788 AA; 83926 MW;
SEQUENCE	788 AA;	Query Match	21.3%; Score 509.5; DB 17; Length 789;
Best Local Similarity	30.65%; Pred. No. 3.9e-14;	Best Local Similarity	21.2%; Score 507.5; DB 1; Length 788;
Matches	144; Conservative 80; Mismatches 165; Indels 81; Gaps 8;	Matches	30.9%; Pred. No. 4.8e-14;
QY	21 ALADRICLDEARIAWRSLFTGIDDITMAALAAEQPLPEATA DALAVIDFYDILESTERTQD 80	QY	21 ALADRICLDEARIAWRSLFTGIDDITMAALAAEQPLPEATA DALAVIDFYDILESTERTQD 80
Db	361 SLQDRI-QEAE-----A VEEANAKAAKEAEELRDAEAQAEERAK- 401	Db	361 SLQDRI-QEAE-----A VEEANAKAAKEAEELRDAEAQAEERAK- 401
QY	81 LFANSKTKTVEQLKETAEYLIGLGRGEYDTEYAQRARIKGKHDLVGLGPDVYLGYATRY 140	QY	81 LFANSKTKTVEQLKETAEYLIGLGRGEYDTEYAQRARIKGKHDLVGLGPDVYLGYATRY 140
Db	402 -ATAAAESRLQERAAD-----417	Db	402 -ATAAAESRLQERAAD-----417
QY	141 YTGLLDALADDVADGEAEA--AAVDELVARFLPMKLTFDQIAMDYIDSYAQRLH 198	QY	141 YTGLLDALADDVADGEAEA--AAVDELVARFLPMKLTFDQIAMDYIDSYAQRLH 198
Db	418 YSEMVQAVAGDLTERLDEDADEBAAVATEFNAML----DGLEBATIAQVAGADEVA 472	Db	418 YSEMVQAVAGDLTERLDEDADEBAAVATEFNAML----DGLEBATIAQVAGADEVA 472
QY	199 DEIDSROQELANAVATHVEAPLSSLEATSQDAERTDTMRARTDDQYDRMDSREISSVS 258	QY	199 DEIDSROQELANAVATHVEAPLSSLEATSQDAERTDTMRARTDDQYDRMDSREISSVS 258
Db	473 DE-----TLQVATGAE---EIELTSQYTSERIQLAQGEAAADDALATMTIDEATDGVTAQGEOLGERAA 318	Db	473 DE-----TLQVATGAE---EIELTSQYTSERIQLAQGEAAADDALATMTIDEATDGVTAQGEOLGERAA 318
QY	259 ASVEEVASTADDVVRTSEDAA PALAQGEAAADDALATMTIDEATDGVTAQGEOLGERAA 318	QY	259 ASVEEVASTADDVVRTSEDAA PALAQGEAAADDALATMTIDEATDGVTAQGEOLGERAA 318
Db	522 ASIQEVAAASSVATAADAVERGEAGRDAAESADDMATEESLSADAQDILALQERMS 581	Db	522 ASIQEVAAASSVATAADAVERGEAGRDAAESADDMATEESLSADAQDILALQERMS 581
QY	319 DVESTGVDDIAETTFTVQLDEVNORIGEYVEREEAMETLQETDAVEDAASMQEVSTATDE 438	QY	319 DVESTGVDDIAETTFTVQLDEVNORIGEYVEREEAMETLQETDAVEDAASMQEVSTATDE 438
Db	582 DIGDIEFIFTIAEQTNMLAINAN EAARDQGDFAVANEVEQDAAETIKQAADIES 641	Db	582 DIGDIEFIFTIAEQTNMLAINAN EAARDQGDFAVANEVEQDAAETIKQAADIES 641
QY	379 LVEQMQAETETVQLDEVNORIGEYVEREEAMETLQETDAVEDAASMQEVSTATDE 438	QY	379 LVEQMQAETETVQLDEVNORIGEYVEREEAMETLQETDAVEDAASMQEVSTATDE 438
Db	642 EIQAQAEETDETVAIRATSEHIDCVSTVERAAAEDTYDVAITEDANHICQEISDATED 701	Db	642 EIQAQAEETDETVAIRATSEHIDCVSTVERAAAEDTYDVAITEDANHICQEISDATED 701
QY	439 QAVSTEEVAMNDGVDDRAETIAALDDIADATDOQVTVEEVRETVGKL 488	QY	439 QAVSTEEVAMNDGVDDRAETIAALDDIADATDOQVTVEEVRETVGKL 488

Qy	439	QAVSTTEEVAEMVDGUDRAGEIAAALDDIADATDQVRTVEV	481
Db	702	QDADATQSVVRVVDVDAISQSHVTEAQVSAAAEQSASVAB	744

Search completed: August 10, 2004, 15:28:39
Job time : 60.609 secs

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6	Post-processing: Minimum Match 0% Maximum DB seq length: 0	RESULT 1 ID AAY29391 standard; protein; 153 AA. XX AC AAY29391; XX DT 01-OCT-1999 (first entry) XX DE Sperm whale myoglobin protein sequence. XX KW Myoglobin; Mb; sperm whale; protein design; function; property; XX OS Physeter sp. XX PN JP1193297-A. XX PD 21-JUL-1999. XX PF 06-OCT-1999; XX PR 06-OCT-1999; XX PR 06-OCT-1999; XX PR 06-OCT-1999;
OM protein - protein search, using SW mode!			
Run on: August 10, 2004, 15:18:30 ; Search time 26.4879 Seconds			
Title: US-09-455-978B-76			
Perfect score: 794			
Sequence: 1 VLSEGEWQLVLHWAKEAD.....ALELFRKDIAAKYKELGYQG 153			
Scoring table: BIOSIN62			
Gapext 10.0 , Gapext 0.5			
Searched: 1586107 seqs, 282547305 residues			
Total number of hits satisfying chosen parameters: 1586107			
Minimum DB seq length: 0			
Maximum DB seq length: 20000000000			
Post-processing: Minimum Match 0% Listing first 45 summaries			
Database : A_Geneseq_29Jan04:*			
1: geneseqp1980s:*			
2: geneseqp1930s:*			
3: geneseqp2000s:*			
4: geneseqp2001s:*			
5: geneseqp2002s:*			
6: geneseqp2003as:*			
7: geneseqp2003bs:*			
8: geneseqp2004s:*			
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.			
SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	794	100.0 153 AAY29391	Aay29391 Sperm wha
2	794	100.0 153 2 AW81769	Aay81769 Whale MYG
3	794	100.0 153 3 AAY69975	Aay69975 MYGL prot
4	715	90.1 154 2 AW62270	Aaw62270 Modified
5	713	89.8 154 2 AW29741	Aaw29741 Modified
6	711	89.5 154 2 AW29739	Aaw29739 Modified
7	711	89.5 154 2 AW29740	Aaw29740 Modified
8	711	89.5 154 2 AW62271	Aaw62271 Modified
9	707	89.0 154 2 AW29742	Aaw29742 Modified
10	701	88.3 154 2 AW29743	Aaw29743 Modified
11	700	88.2 154 2 AW62272	Aaw62272 Modified
12	699	88.0 154 2 AW29744	Aaw29744 Modified
13	699	88.0 166 7 ADC31105	Adc31106 Human nov
14	699	88.0 206 7 ADC32817	Aad32817 Human nov
15	697	87.8 154 2 AW62273	Aaw62273 Modified
16	613	77.2 202 4 ARG04708	Abg04708 Novel hum
17	535	67.4 231 4 ARG04706	Abg04706 Novel hum
18	472	59.4 113 3 AG00022	Aag00022 Human sec
19	386	48.7 334 4 AGC21418	Abg21418 Novel hum
20	314	39.5 71 3 AG00021	Aag00021 Human sec
21	253	31.9 107 4 ABG21417	Abg21417 Novel hum
22	243	30.6 64 2 ARG5437	Aarg5437 Apomyoglo
23	240	30.2 60 2 AAR95436	Aar95436 Apomyoglo
24	215.5	27.1 190 5 ABB07967	Abb07967 Rat STUF
25	215.5	27.1 201 2 AAW86334	Aaw86334 Kidney in

ALIGNMENTS

Aae04661 Haem prot
Aar95435 Apomyoglo
Aab07966 Human STU
Adc10192 Human NOV
Adc10194 Human NOV
Aab04662 Haem prot
Aae04665 Haem prot
Aae04664 Haem prot
Aae04669 Haem prot
Aae04670 Haem prot
Aae04673 Haem prot
Aae04672 Haem prot
Aae04663 Haem prot
Aay29393 Sperm wha
Aae04671 Haem prot
Aab41929 Human ORF
Aab04674 Haem prot
Abg04707 Novel hum
Abg04705 Novel hum
Aaw73385 Human hae

The present invention describes a method for producing an ideal amino acid sequence. The method comprises: (1) preparation of one initial amino acid sequence corresponding to the structure of a protein; (2) selecting the amino acid residues optimal to each site; (3) repeating the steps of selecting the amino acid residues optimum to each site of the total amino acid residues constituting the N-order amino acid sequence (N is an integer not less than 2), to give an N-1-order amino acid sequence, consisting of the selected amino acid residues until the N-order amino acid sequence comes to be same as the N+1-order amino acid sequence; and (4) selecting the resultant N-order amino acid sequence as the optimum amino acid sequence of said protein. The method can design a protein with desired functions and properties. The present sequence represents sperm whale myoglobin used in an example from the present invention where the stereo structure of sperm whale myoglobin was targeted as a globular protein to detect an amino acid sequence best fit to the structure of the main chain of the sperm whale myoglobin

Db	61 LKKHGVTVLGALKKGHHEAEUKPQASHATKHKPIKYLEFISAAIIHVLSRHP 120	Db	3 LSDGEWQQVILNVNGKVEADIAHGQEVLRILFTGHPETLEKFDFEKKLKEAEMKAEDL 62
Qy	121 GDFGADAQGAMMNKAELFRKDIAAKYKELGQG 153	Qy	62 KKHGTIVTALGALLRKKGHEAELKPLAQSHATKHKPIKYLEFISAAIIHVLSRHPG 121
Db	121 GDFGADAQGAMMNKAELFRKDIAAKYKELGQG 153	Db	63 KKHGTIVTALGGILKKKGHEAELKPLAQSHATKHKPIKYLEFISAAIIHVLSRHPG 122
		Qy	122 DFGADAGAMMNKAELFRKDIAAKYKELGQG 153
		Db	123 DFGADAGAMMNKAELFRKDIAAKYKELGQG 154
RESULT 4		RESULT 5	
ID AAW62270	standard; protein; 154 AA.	ID AAW29741	AAW29741 standard; protein; 154 AA.
XX		XX	
AC AAW62270;		AC AAW29741;	
XX		XX	
DT 26-OCT-1998	(first entry)	DT 26-OCT-1998	(first entry)
XX		XX	
DE Modified myoglobin protein 1.		DE Modified myoglobin protein 7.	
XX		XX	
KW Heart; myoglobin; Mb; random mutagenesis; Peroxidase activity; Fe(III); Mn(II) Peroxidase activity; hydrogen peroxide; Mn binding site; catalyse; phenolic; polymer; antigen; antibody; diagnostic assay.		KW Heart; myoglobin; Mb; random mutagenesis; peroxidase activity; Fe(III); Mn(II) peroxidase activity; hydrogen peroxide; Mn binding site; catalyse; phenolic; polymer; antigen; antibody; diagnostic assay.	
XX		XX	
OS Equus sp.		OS Equus sp.	
OS Synthetic.		OS Synthetic.	
XX		XX	
CA2185605-A.		CA2185605-A.	
PN		PN	CA2185605-A.
XX		XX	
PD 17-MAR-1998.		PD 17-MAR-1998.	
XX		XX	
PF 16-SEP-1996;	96CA-02185605.	PF 16-SEP-1996;	96CA-02185605.
XX		XX	
PR 16-SEP-1996;	96CA-02185605.	PR 16-SEP-1996;	96CA-02185605.
XX		XX	
PA (MAUK/)	MAUK A. G.	PA (MAUK/)	MAUK A. G.
PA (WANL/)	WAN L.	PA (WANL/)	WAN L.
PA (LEEH/)	LEE H.	PA (LEEH/)	LEE H.
PA (BRAY/)	BRAYER G. D.	PA (BRAY/)	BRAYER G. D.
PA (TONG/)	TONG H.	PA (TONG/)	TONG H.
PA (SMIT/)	SMITH M.	PA (SMIT/)	SMITH M.
XX		XX	
PI Wan L., Mauk AG, Lee H, Brayer GD, Tong H, Smith M;		PI Wan L., Mauk AG, Lee H, Brayer GD, Tong H, Smith M;	
WPI; 1998-388655/34.		WPI; 1998-388655/34.	
XX		XX	
PT New modified myoglobin with amino acid alterations has increased peroxidase activity - used for catalysing oxidation of substrates with peroxides, e.g. in waste treatment and as label for antigens.		PT New modified myoglobin with amino acid alterations has increased peroxidase activity - used for catalysing oxidation of substrates with peroxides, e.g. in waste treatment and as label for antigens.	
XX		XX	
PT PT sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic horse heart myoglobin (Mb) proteins, produced by random mutagenesis of the synthetic Mb gene. This present sequence contains a substitution at residue 45 which has been shown to enhance the peroxidase activity of the Mb protein in general, and it may also exhibit enhanced Mn(II) peroxidase activity as well. This is due to the enhanced rate of reaction of Fe(III) Mb with hydrogen peroxide and also because the substitution residue 45 creates a new Mn binding site which is stronger than the natural Mn binding site of Mb. These protein can be used to catalyse oxidation of a substrate by peroxidase, e.g. oxidation of Mn2+. Peroxidases are used for oxidising phenolics to polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in diagnostic assays and for oxidative treatment of waste streams. NB. This sequence is not given in the specification but was created from the wild type sequence given by the inventors. The position of the substitution given does not include the initiation codon		PT PT sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic horse heart myoglobin (Mb) proteins, produced by random mutagenesis of the synthetic Mb gene. This present sequence contains a substitution at residue 45 which has been shown to enhance the peroxidase activity of the Mb protein in general, and it may also exhibit enhanced Mn(II) peroxidase activity as well. This is due to the enhanced rate of reaction of Fe(III) Mb with hydrogen peroxide and also because the substitution residue 45 creates a new Mn binding site which is stronger than the natural Mn binding site of Mb. These protein can be used to catalyse oxidation of a substrate by peroxidase, e.g. oxidation of Mn2+. Peroxidases are used for oxidising phenolics to polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in diagnostic assays and for oxidative treatment of waste streams. NB. This sequence is not given in the specification but was created from the wild type sequence given by the inventors. The position of the substitution given does not include the initiation codon	
XX		XX	
PS Claim 15; Page; 23pp; English.		PS Claim 10; Page; 23pp; English.	
XX		XX	
CC Sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic horse heart myoglobin (Mb) proteins, produced by random mutagenesis of the synthetic Mb gene. This present sequence contains a substitution at residue 45 which has been shown to enhance the peroxidase activity of the Mb protein in general, and it may also exhibit enhanced Mn(II) peroxidase activity as well. This is due to the enhanced rate of reaction of Fe(III) Mb with hydrogen peroxide and also because the substitution residue 45 creates a new Mn binding site which is stronger than the natural Mn binding site of Mb. These protein can be used to catalyse oxidation of a substrate by peroxidase, e.g. oxidation of Mn2+. Peroxidases are used for oxidising phenolics to polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in diagnostic assays and for oxidative treatment of waste streams. NB. This sequence is not given in the specification but was created from the wild type sequence given by the inventors. The position of the substitution given does not include the initiation codon		CC Sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic horse heart myoglobin (Mb) proteins, produced by random mutagenesis of the synthetic Mb gene. This present sequence contains a substitution at residue 45 which has been shown to enhance the peroxidase activity of the Mb protein in general, and it may also exhibit enhanced Mn(II) peroxidase activity as well. This is due to the enhanced rate of reaction of Fe(III) Mb with hydrogen peroxide and also because the substitution residue 45 creates a new Mn binding site which is stronger than the natural Mn binding site of Mb. These protein can be used to catalyse oxidation of a substrate by peroxidase, e.g. oxidation of Mn2+. Peroxidases are used for oxidising phenolics to polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in diagnostic assays and for oxidative treatment of waste streams. NB. This sequence is not given in the specification but was created from the wild type sequence given by the inventors. The position of the substitution given does not include the initiation codon	
XX		XX	
PS Sequence 154 AA;		PS Sequence 154 AA;	
XX		XX	
Query Match 90.1%; Score 715; DB 2; Length 154;		Query Match 89.8%; Score 713; DB 2; Length 154;	
Best Local Similarity 88.2%; Pred. No. 2e-71;		Best Local Similarity 87.5%; Pred. No. 3.3e-7;	
Matches 134; Conservative 8; Mismatches 10; Indels 0; Gaps 0;		Matches 133; Conservative 9; Mismatches 10; Indels 0; Gaps 0;	
Qy 2 ISEGEGWQLVLIHWAKYEADIAHGQEVLRILFKSHPTLEKFDRKHLKTEAEMKAEDL 61		Qy 2 ISEGEGWQLVLIHWAKYEADIAHGQEVLRILFKSHPTLEKFDRKHLKTEAEMKAEDL 61	

Db	3 LSDGEWQQLNVGKVYEAIDIAGHGQEVLIRLFTGHPETLEKDPDFKHLKTEAEMKASEDL 62	Db	3 LSDGEWQQLNVGKVYEAIDIAGHGQEVLIRLFTGHPETLEKDPDFKHLKTEAEMKASEDL 62
Qy	62 KKHGTIVTIALGILKKHHAEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPG 121	Qy	62 KKHGTIVTIALGILKKHHAEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPG 121
Db	63 KKHGTIVTIALGILKKHHAEALKPLAQSHATKHKPIKYLEFSSDAIHVVLHSRHPG 122	Db	63 KKHGTIVTIALGILKKHHAEALKPLAQSHATKHKPIKYLEFISDAIHVVLHSRHPG 122
Qy	122 DFGADAGQAMMKALELPFDKTDIAAKYKELGYQG 153	Qy	122 DFGADAGQAMMKALELPFDKTDIAAKYKELGYQG 153
Db	123 DFGADAGQAMMKALELPFDKTDIAAKYKELGFQG 154	Db	123 DFGADAGQAMMKALELPFDKTDIAAKYKELGFQG 154
<hr/>			
RESULT 6			
AAW29739	AAW29739 standard; protein; 154 AA.	AAW29740	AAW29740 standard; protein; 154 AA.
XX		XX	
AC	AAW29739;	AC	AAW29740;
XX		XX	
DR	26-OCT-1998 (first entry)	DT	26-OCT-1998 (first entry)
XX		XX	
DE	Modified myoglobin protein 5.	DE	Modified myoglobin protein 6.
XX		XX	
XX	Heart; myoglobin; Mb; random mutagenesis; Peroxidase activity; Fe(III);	XX	Heart; myoglobin; Mb; random mutagenesis; Peroxidase activity; Fe(III);
KW	Mn(II) peroxidase activity; hydrogen peroxide; Mn binding site; catalyse;	KW	Mn(II) peroxidase activity; hydrogen peroxide; Mn binding site; catalyse;
KW	phenolic; polymer; antigen; antibody; diagnostic assay.	KW	phenolic; polymer; antigen; antibody; diagnostic assay.
XX		XX	
OS	Equus sp.	OS	Equus sp.
OS	Synthetic.	OS	Synthetic.
XX		XX	
PN	CA2185605-A.	PN	CA2185605-A.
XX		XX	
PD	17-MAR-1998.	PD	17-MAR-1998.
XX		XX	
PF	16-SEP-1996; 96CA-02185605.	PF	16-SEP-1996; 96CA-02185605.
XX		XX	
PR	16-SEP-1996; 96CA-02185605.	PR	16-SEP-1996; 96CA-02185605.
XX		XX	
PA	(MAUK/) MAUK A. G.	PA	(MAUK/) MAUK A. G.
PA	(WANL/) WAN L.	PA	(WANL/) WAN L.
PA	(LEEH/) LEE H.	PA	(LEEH/) LEE H.
PA	(BRAY/) BRAYER G. D.	PA	(BRAY/) BRAYER G. D.
PA	(TONG/) TONG H.	PA	(TONG/) TONG H.
PA	(SMIT/) SMITH M.	PA	(SMIT/) SMITH M.
XX		XX	
PI	Wan L, Mauk AG, Lee H, Brayer GD, Tong H, Smith M;	PI	Wan L, Mauk AG, Lee H, Brayer GD, Tong H, Smith M;
XX		XX	
DR	WPI; 1998-388655/34.	DR	WPI; 1998-388655/34.
XX		XX	
PT	New modified myoglobin with amino acid alterations has increased peroxidase activity - used for catalysing oxidation of substrates with peroxides, e.g. in waste treatment and as label for antigens.	PT	New modified myoglobin with amino acid alterations has increased peroxidase activity - used for catalysing oxidation of substrates with peroxides, e.g. in waste treatment and as label for antigens.
XX		XX	
PS	Claim 10; Page: 23pp; English.	PS	Claim 10; Page: 23pp; English.
XX		XX	
CC	Sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic horse heart myoglobin (Mb) proteins, produced by random mutagenesis of the synthetic Mb gene. This present sequence contains a substitution at Thr91Le which has been shown to enhance the peroxidase activity of Mb. These protein can be used to catalyse oxidation of a substrate by peroxidase, e.g. oxidation of Mn ²⁺ . Peroxidases are used for oxidising phenolics to polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in diagnostic assays and for oxidative treatment of waste streams. NB. This sequence is not given in the specification but was created from the wild type sequence given by the inventors. The position of the substitution given does not include the initiation codon	CC	Sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic horse heart myoglobin (Mb) proteins, produced by random mutagenesis of the synthetic Mb gene. This present sequence contains a substitution at Thr91Le which has been shown to enhance the peroxidase activity of Mb. These protein can be used to catalyse oxidation of a substrate by peroxidase, e.g. oxidation of Mn ²⁺ . Peroxidases are used for oxidising phenolics to polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in diagnostic assays and for oxidative treatment of waste streams. NB. This sequence is not given in the specification but was created from the wild type sequence given by the inventors. The position of the substitution given does not include the initiation codon
CC		CC	
CC	Sequence 154 AA;	CC	Sequence 154 AA;
CC		CC	
Qy	Query Match 89.5%; Score 711; DB 2; Length 154; Best Local Similarity 87.5%; Pred. No. 5.5e-71; Matches 133; Conservative 9; Mismatches 10; Indels 0; Gaps 0;	Qy	Query Match 89.5%; Score 711; DB 2; Length 154; Best Local Similarity 87.5%; Pred. No. 5.5e-71; Matches 133; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
Qy	2 LSEGENMOLVLYWAKYEVADAVGHGQDILRLFKSHPBTLEKDFRKLKTEAEMKASEDL 61	Qy	2 LSEGENMOLVLYWAKYEVADAVGHGQDILRLFKSHPBTLEKDFRKLKTEAEMKASEDL 61

Db 3 LSDGEWQQVLNWVGKVEADIAGHQEVILRLFTGHPETEKFDKLKHLEAEMKAEDL 62
 Qv 62 KKHGVTVLTLGAIKKKGHEAEKLPLAQSHATKHKPIKYLERFISMAIINVLSRHPG 121
 Db 63 KRHGTVVTLALGGTLKKGKGHEAEKLPLAQSHATKHKPIKYLEFISDAIIHVLSRHPG 122
 Db 122 DFGADAOQAMMKALEFLRKDIIAKYKEGYQG 153
 Qv 123 DFGADAOQAMMTKALEFLRKDIIAKYKEGFQG 154

RESULT 8
 AAW62271 ID AAW62271 standard; protein; 154 AA.
 XX AC AAW62271:
 XX DT 26-OCT-1998 (first entry)
 XX DE Modified myoglobin protein 2.
 XX KW Heart; myoglobin; Mb; random mutagenesis; peroxidase activity; Fe(III);
 Mn(II); peroxidase activity; hydrogen peroxide; Mn binding site; catalyse;
 KW phenolic; polymer; antigen; antibody; diagnostic assay.
 XX OS Equus sp.
 OS Synthetic.
 XX CA2185605-A.
 XX PR 17-MAR-1998.
 XX PF 16-SEP-1996; 96CA-02185605.
 XX PR 16-SEP-1996; 96CA-02185605.
 XX PA (MAUK/) MAUK A. G.
 PA (WANL/) WAN L.
 PA (LEEH/) LEE H.
 PA (BRAY/) BRAYER G. D.
 PA (TONG/) TONG H.
 PA (SMIT/) SMITH M.
 XX PI Wan L, Mauk AG, Lee H, Brayer GD, Tong H, Smith M,
 DR WPI: 1998-388655/34.
 XX New modified myoglobin with amino acid alterations has increased
 peroxidase activity - used for catalysing oxidation of substrates with
 peroxides, e.g. in waste treatment and as label for antigens.
 XX PS Claim 6; Page: 23pp; English.

CC Sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic
 horse heart myoglobin (Mb) proteins, produced by random mutagenesis of
 the synthetic Mb gene. This present sequence contains a substitution at
 Lys45Glu, which has been found to enhance the peroxidase activity of the
 protein in general, and it may also exhibit enhanced Mn(III) peroxidase
 activity as well. This is due to the enhanced rate of reaction of Fe(III)
 Mb with hydrogen peroxide and also because the substitution at residue 45
 creates a new Mn binding site which is stronger than the natural Mn
 binding site of Mb. This protein also contains a substitution at
 Lys63Glu, which was also found to be required for Mn binding. These
 protein can be used to catalyse oxidation of a substrate by peroxidase,
 e.g. oxidation of Mn²⁺. Peroxidases are used for oxidising phenolics to
 polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in
 diagnostic assays and for oxidative treatment of waste streams. NB. This
 sequence is not given in the specification but was created from the wild
 type sequence given by the inventors. The positions of the substitutions
 given do not include the initiation codon
 CC Sequence 154 AA;
 SQ

Query Match 89.5%; Score: 711; DB 2; Length: 154;
 Best Local Similarity 87.5%; Pred. No. 5.5e-71;
 Matches 133; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Qv 2 LSGEWQVLNWVGKVEADIAGHQEVILRLFTGHPETEKFDFKLKHLEAEMKAEDL 61
 Db 3 LSDGEWQQVLNWVGKVEADIAGHQEVILRLFTGHPETEKFDFKLKHLEAEMKAEDL 62
 Qv 62 KRHGTVVTLALGGTLKKGKGHEAEKLPLAQSHATKHKPIKYLEFISDAIIHVLSRHPG 121
 Db 63 KRHGTVVTLALGGTLKKGKGHEAEKLPLAQSHATKHKPIKYLEFISDAIIHVLSRHPG 122
 Qv 122 DFGADAOQAMMKALEFLRKDIIAKYKEGYQG 153
 Db 123 DFGADAOQAMMTKALEFLRKDIIAKYKEGFQG 154

RESULT 9
 AAW29742 ID AAW29742 standard; protein; 154 AA.
 XX AC AAW29742:
 XX DT 26-OCT-1998 (first entry)
 XX DE Modified myoglobin protein 8.
 XX KW Heart; myoglobin; Mb; random mutagenesis; peroxidase activity; Fe(III);
 Mn(II); peroxidase activity; hydrogen peroxide; Mn binding site; catalyse;
 KW phenolic; polymer; antigen; antibody; diagnostic assay.
 XX OS Equus sp.
 OS Synthetic.
 XX PN CA2185605-A.
 XX PD 17-MAR-1998.
 XX PF 16-SEP-1996; 96CA-02185605.
 XX PR 16-SEP-1996; 96CA-02185605.
 XX PA (MAUK/) MAUK A. G.
 PA (WANL/) WAN L.
 PA (LEEH/) LEE H.
 PA (BRAY/) BRAYER G. D.
 PA (TONG/) TONG H.
 PA (SMIT/) SMITH M.
 XX PI Wan L, Mauk AG, Lee H, Brayer GD, Tong H, Smith M,
 DR WPI: 1998-388655/34.
 XX New modified myoglobin with amino acid alterations has increased
 peroxidase activity - used for catalysing oxidation of substrates with
 peroxides, e.g. in waste treatment and as label for antigens.
 XX PS Claim 10; Page: 23pp; English.

CC Sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic
 horse heart myoglobin (Mb) proteins, produced by random mutagenesis of
 the synthetic Mb gene. This present sequence contains substitutions with
 the synthetic Mb gene. This present sequence contains substitutions at
 Phe41Leu and Ile107Phe which have been shown to enhance the peroxidase
 activity of Mb. These protein can be used to catalyse oxidation of a
 substrate by peroxidase, e.g. oxidation of Mn²⁺. Peroxidases are used for
 oxidising phenolics to polymers, as labels for e.g. antigens, antibodies
 or oligonucleotides, in diagnostic assays and for oxidative treatment of
 waste streams. NB. This sequence is not given in the specification but
 was created from the wild type sequence given by the inventors. The initiation codon
 CC positions of the substitutions given do not include the initiation codon
 XX Sequence 154 AA;
 SQ

RESULT 10
AAW29743 standard; protein; 154 AA.
XX

RESULT 11
AAW62272
ID AAW62272 standard; protein; 154 AA.

XX	26-OCT-1998	(first entry)
XX	Modified myoglobin protein 9.	
XX	Heart; myoglobin; Mb; random mutagenesis; peroxidase activity; Fe(III); Mn(II) Peroxidase activity; hydrogen peroxide; Mn binding site; catalyse; phenolic; polymer; antigen; antibody; diagnostic assay.	
XX		
AC		Hawazzi;
XX	DT	26-OCT-1998 (first entry)
XX	DE	Modified myoglobin protein 3.
XX	KW	Heart; myoglobin; Mb; random mutagenesis; peroxidase activity; Fe(III); Mn(II) peroxidase activity; hydrogen peroxide; Mn binding site; catalyse; phenolic; polymer; antigen; antibody; diagnostic assay.
XX	KW	
XX	KW	
XX	KW	

OS	OS	equus sp.
Synthetic.	Synthetic.	
XX	XX	
CA2185605-A.	CA2185605-A.	
PN	PN	
CA2185605-XX	CA2185605-XX	
IPD	IPD	
17-MAR-1998.	17-MAR-1998.	
XX	XX	
PPF	PPF	
16-SEP-1996;	96CA-02185605.	
XXX	XXX	
PR	PR	
16-SEP-1996;	96CA-02185605.	
XX	XX	

PA	(WANL/)	WAN L.						
PA	(LEEH/)	LEE H.						
PA	(BRAY/)	BRAYER G. D.						
PA	(TONG/)	TONG H.						
PA	(SMIT/)	SMITH M.						
XXX								
PI	Wan L,	Mauk AG,	Lee H,	Brayer GD,	Tong H,	Smith M;		
PI								
PI								
XX								
XX								
DR								
WPI;	1998-388555/34.							

New modified myoglobin with amino acid alterations has increased peroxidase activity - used for catalysing oxidation of substrates with peroxides, e.g. in waste treatment and as label for antigens.

CC polymers, as labels for e.g. antigens, antibodies or oligonucleotides, in CC diagnostic assays and for oxidative treatment of waste streams. NB. This CC sequence is not given in the specification but was created from the wild type sequence given by the inventors. The positions of the substitutions CC given do not include the initiation codon

XX Sequence 154 AA;

Query Match 88.2%; Score 700; DB 2; Length 154;

Best Local Similarity 86.8%; Pred. No. 9.3e-70;
Matches 132; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Db 2 LSEGEWQVILHYAKVEADAVAGHQDILIRLJKSHPETLEKDFRKHLKTEAMKASEDL 61

Db 3 LSGEWWQVILHYAKVEADAVAGHQDILIRLJKSHPETLEKDFRKHLKTEAMKASEDL 62

Qy 62 KKHGVTVLTAQGALKKKGGHHEELPLAQSHATKHKIPKYLEFISAAITHVLHSRHPG 121

Db 63 KHBGTVVLTALG3LKKKGHHAEELPLAQSHATKHKIPKYLEFISAAITHVLHSRHPG 122

Qy 122 DFGADAQGAMNKALELFRKDIAAKYKELGYQG 153

Db 123 DFGADAQGAMNKALELFRKDIAAKYKELGFQG 154

RESULT 12 AAW29744

ID AAW29744 standard; protein; 154 AA.

XX AC AAW29744;

XX DT 26-OCT-1998 (first entry)

XX XX Modified myoglobin protein 10.

XX Heart; myoglobin; Mb; random mutagenesis; peroxidase activity; Fe(III);

KW Mn(II) peroxidase activity; hydrogen peroxide; Mn binding site; catalyse;

KW phenolic; polymer; antibody; diagnostic assay.

XX XX Equus sp.

XX Synthetic.

XX PN CA2185605-A.

XX PD 17-MAR-1998.

XX 16-SEP-1996; 96CA-02185605.

XX 16-SEP-1996; 96CA-02185605.

PR (TONG/) TONG H.

PA (MAUK/) MAUK A. G.

PA (WANL/) WAN L.

PA (LEEH/) LEE H.

PA (BRAY/) BRAYER G. D.

PA (SMIT/) SMITH M.

XX XX Wan L., Mauk AG., Lee H., Brayer GD., Tong H., Smith M;

XX DR; 1998-388655/34.

XX PT New modified myoglobin with amino acid alterations has increased peroxidase activity - used for catalysing oxidation of substrates with

PT peroxides, e.g. in waste treatment and as label for antigens.

Claim 21; Page: 23pp; English.

XX Sequences AAW62270-W62273, and AAW29739-W29744, are modified synthetic

CC horse heart myoglobin (Mb) proteins, produced by random mutagenesis of

CC the synthetic Mb gene. This present sequence contains a substitution at

CC Lys4Glu, which has been found to enhance the peroxidase activity of the

CC protein in general, and it may also exhibit enhanced Mn(II) peroxide

CC activity as well. This is due to the enhanced rate of reaction of Fe(III)

CC Mb with hydrogen peroxide and also because the substitution at residue 45 CC creates a new Mn binding site which is stronger than the natural Mn CC binding site of Mb. This sequence also contains substitutions at the CC Thr39Ile, Phe46Leu, and Ile107Phe, which have been shown to enhance the CC peroxidase activity of Mb. These protein can be used to catalyse oxidation CC of a substrate by peroxidase, e.g. oxidation of Mn²⁺. Peroxidases are CC used for oxidising phenolics to polymers, as labels for e.g. antigens, CC antibodies or oligonucleotides, in diagnostic assays and for oxidative CC treatment of waste streams. NB. This sequence is not given in the CC specification but was created from the wild type sequence given by the CC inventors. The positions of the substitutions given do not include the CC initiation codon

XX SQ Sequence 154 AA;

Query Match 88.0%; Score 699; DB 2; Length 154;

Best Local Similarity 86.2%; Pred. No. 1.2e-69; Matches 131; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Db 2 LSEGEWQVILHYAKVEADAVAGHQDILIRLJKSHPETLEKDFRKHLKTEAMKASEDL 61

Db 3 LSDGEWWQVILHYAKVEADAVAGHQDILIRLJKSHPETLEKDFRKHLKTEAMKASEDL 62

Qy 62 KKHGVTVLTAQGALKKKGGHHEELPLAQSHATKHKIPKYLEFISAAITHVLHSRHPG 121

Db 63 KKHGTVVLTALG3LKKKGHHAEELPLAQSHATKHKIPKYLEFISAAITHVLHSRHPG 122

Qy 122 DFGADAQGAMNKALELFRKDIAAKYKELGYQG 153

Db 123 DFGADAQGAMNKALELFRKDIAAKYKELGFQG 154

RESULT 13 AAW29744

ID AAW29744 standard; protein; 154 AA.

XX AC AAW29744;

XX DT 18-DEC-2003 (first entry)

XX XX Human novel polypeptide sequence, SEQ ID NO:1188.

XX ID ADC31106 standard; protein; 166 AA.

XX AC ADC31106;

XX DT 18-DEC-2003 (first entry)

XX XX Human novel polypeptide sequence, SEQ ID NO:1188.

XX DE Human; diagnostic; drug screening; forensics; gene mapping;

XX KW Alzheimer's disease; Alzheimer's disease; Parkinson's disease;

XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

XX KW ulcers; osteoporosis; autoimmune disease; cancer;

XX KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

XX KW neuroprotective; antiinflammatory; thrombolytic; vulnerary;

XX KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

XX KW gene therapy; chromosome 22q13.1.

XX OS Homo sapiens.

XX PN WO2003029271-A2.

XX PD 10-APR-2003.

XX PP 24-SEP-2002; 2002WO-US03474.

XX PR 24-SEP-2001; 2001US-0324631P.

XX PA (HYSEQ INC.

XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Drmanac RT;

PI Haley-Vicente D, Drmanac RT;

XX DR WPI; 2003-371981/35.

XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or PT treating conditions such as neurodegenerative diseases, anemias, platelet PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or PT

PT cancer.
 XX Claim 20; SEQ ID NO 1188; 1185PP; English.
 PS The invention relates to 971 novel human cDNA sequences (ADC29919-
 XX ADC30899) and the polypeptides they encode (ADC10890-ADC1860). The
 CC invention also relates to nucleic acid sequences over 90% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contiguous sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627); and the polypeptides encoded by the contigs
 CC (ADC33294). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The Sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 166 AA:
 SQ Query Match 88.0%; Score 699; DB 7; Length 166;
 Best Local Similarity 84.9%; Pred. No. 1.3e-69;
 Matches 129; Conservative 13; Mismatches 10; Indels 0; Gaps 0;
 OY 2 LSEGEGWQLVHLWAKVEADVAGHQDILIRLFLKSHPTELEKFDRFKHLKTEAMKASDDL 61
 Db 15 LSDGEWQQLVLNWKGKEAIPGHEQEVLIPRLFKSHPTELEKFDRFKHLKSEDEMKASDL 74
 QY 62 KKHGATVTALGAIKKKGHEARLKPLAQSHATKHPKPIKYLEFISBAILHVHLHSRHPG 121
 Db 75 KKHGATVTALGGTLLKRGHHEAIEIKPLAQSHATKHPKPVKYLEFISBCTIQVLQSKHPG 134
 QY 122 DFGADAQGMNKKALELFRDIAAKYKELGYQG 153
 Db 135 DFGADAQGMNKKALELFRDMSNYKELGFQG 166

RESULT 14
 ADC32817 Human novel contig-encoded polypeptide sequence, SEQ ID NO:2899.
 ID ADC32817 standard; protein; 206 AA.
 AC ADC32817;
 XX DT 18-DEC-2003 (first entry)
 DE Human; diagnostic; drug screening; forensics; gene mapping;
 DE biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 DE neurodegenerative disease; anaemia; platelet disorder; wound; burns;
 DE ulcers; osteoporosis; autoimmune disease; cancer;
 DE molecular weight marker; food supplement; anti-parkinsonian; nootropic;
 DE neuroprotective; anti-nausemic; thrombolytic; vasoconstrictor;
 DE antiulcer; osteopathic; immunosuppressive; anti-inflammatory;
 DE cytostatic;

XX gene therapy; chromosome 22q13.1.

XX Homo sapiens.
 XX WO2003029271-A2.
 PN PD 10-APR-2003.
 XX 24-SEP-2002; 2002WO-US0303474.
 PR 24-SEP-2001; 2001US-03224631P.
 PA (HYSE-) HYSSQ INC.
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T,
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX WPI; 2003-371981/35.
 DR N-FSDB; ADC32050.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemas, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX PS SEQ ID NO 2899; 1185PP; English.
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30899) and the polypeptides they encode (ADC10890-ADC1860). The
 CC invention also relates to nucleic acid sequences over 90% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contiguous sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627); and the polypeptides encoded by the contigs
 CC (ADC33294). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a human contig-
 CC encoded polypeptide sequence used in an example of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 206 AA:
 SQ Query Match 68.0%; Score 699; DB 7; Length 206;
 Best Local Similarity 84.9%; Pred. No. 1.8e-69;
 Matches 129; Conservative 13; Mismatches 10; Indels 0; Gaps 0;
 QY 2 LSEGEGWQLVHLWAKVEADVAGHQDILIRLFLKSHPTELEKFDRFKHLKTEAMKASEDL 61
 Db 55 LSDGEWQQLVLNWKGKEAIPGHEQEVLIPRLFKSHPTELEKFDRFKHLKSEDEMKASDL 114
 QY 62 KKHGATVTALGAIKKKGHEARLKPLAQSHATKHPKPIKYLEFISBAILHVHLHSRHPG 121
 Db 62 KKHGATVTALGGTLLKRGHHEAIEIKPLAQSHATKHPKPVKYLEFISBCTIQVLQSKHPG 134
 DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:2899.
 DE Human; diagnostic; drug screening; forensics; gene mapping;
 DE biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 DE neurodegenerative disease; anaemia; platelet disorder; wound; burns;
 DE ulcers; osteoporosis; autoimmune disease; cancer;
 DE molecular weight marker; food supplement; anti-parkinsonian; nootropic;
 DE neuroprotective; anti-nauseemic; thrombolytic; vasoconstrictor;
 DE antiulcer; osteopathic; immunosuppressive; anti-inflammatory;
 DE cytostatic;

Db 115 KKHGATVLTALGGILKKKGHEAEIKPLAQSHATKHKIPVKYLEFISBCIIQVLQSXHPG 174 Qy 2 LSRGEWQOLVLYHAKVEADVAGHGDIDLIRLFKSHPTELEKDFRKCHLKTEAMKASEDL 61
 Qy 122 DFGADAQGMNKALELFRKDIAAKYKELGYQG 153 Db 3 LSDGEWQOQLNTWGKVEDIAGHQEVILRLPTGHPTLEKDFRKCHLKTEAMKASEDL 62
 Db 175 DFGADAQGMNKALELFRKDMSNYKELGFOQ 206 Qy 62 KRGGVVLTALGAIKKKGHEELKPLAQSHATKHKIPVKYLEFISBCIIQVLHSRHPG 121
 RESULT 15 Db 63 KRGGVVLTALGAIKKKGHEELKPLAQSHATKHKIPVKYLEFISBCIIQVLHSRHPG 122
 AAW62273 ID AAW62273 standard; protein; 154 AA.
 XX Qy 122 DFGADAQGMNKALELFRKDIAAKYKELGYQG 153
 AC DB 123 DFGADAQGMNKALELFRKDMSNYKELGFOQ 154
 XX DT 26-OCT-1998 (first entry)
 XX DB 123 DFGADAQGMNKALELFRKDMSNYKELGFOQ 154
 DE Modified myoglobin protein 4.
 XX Search completed: August 10, 2004, 15:26:14
 KW Job time : 27.4879 secs
 KW Mn(II) peroxidase activity; hydrogen peroxide; Mn binding site; catalyse;
 KW phenolic; Polymer; antigen; antibody; diagnostic assay.
 XX
 OS Equus sp.
 OS Synthetic.
 XX CA2185605-A.
 PN
 PA PD 17-MAR-1998.
 XX PP 16-SEP-1996; 96CA-02185605.
 XX PR 16-SEP-1996; 96CA-02185605.
 XX PA (MAUK/) MAUK A G.
 PA (WANL/) WAN L.
 PA (LEEH/) LEE H.
 PA (BRAY/) BRAYER G D.
 PA (TONG/) TONG H.
 PA (SMIT/) SMITH M.
 XX PI Wan L, Mauk AG, Lee H, Brayer GD, Tong H, Smith M;
 XX DR WPI; 1998-388655/34.
 XX New modified myoglobin with amino acid alterations has increased
 PT peroxidase activity - used for catalysing oxidation of substrates with
 PT peroxides, e.g. in waste treatment and as label for antigens.
 XX Claim 23: Page; 23pp; English.
 XX Sequences AAW62270-WG2273, and AAW29739-W29744, are modified synthetic
 CC horse heart myoglobin (Mb) proteins, produced by random mutagenesis of
 CC the synthetic Mb gene. This present sequence contains a substitution at
 CC Lys45Glu, which has been found to enhance the rate of the
 CC protein in general, and it may also exhibit enhanced Mn(II) peroxidase
 CC activity as well. This is due to the enhanced rate of reaction of Fe(III)
 CC with hydrogen peroxide and also because the substitution at residue 45
 CC creates a new Mn binding site which is stronger than the natural Mn
 CC binding site of Mb. This protein also contains a substitution at
 CC Lys63Glu, which was also found to be required for Mn binding, and
 CC substitutions at His97Leu and Ser92Ala to increase the rate of Mn(II)
 CC turnover. These protein can be used to catalyse oxidation of a substrate
 CC by peroxidase, e.g. oxidation of Mn²⁺. Peroxidases are used for oxidising
 CC phenolics to polymers, as labels for e.g. antigens, antibodies or
 CC oligonucleotides, in diagnostic assays and for oxidative treatment of
 CC waste streams. NB. This sequence is not given in the specification but
 CC was created from the wild type sequence given by the inventors. The
 CC positions of the substitutions given do not include the initiation codon
 XX Sequence 154 AA;
 SQ Query Match 87.8%; Score 697; DB 2; Length 154;
 Best Local Similarity 86.2%; Pred. No. 28-69;
 Matches 131; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

B
I
G
T

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW mode!

Run on: August 10, 2004, 15:23:41 ; Search time 7.77966 Seconds

1015.311 Million cell updates/sec

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	794	100.0	153	3	US-08-932-176-7
2	778.5	98.0	152	3	US-08-932-176-10
3	717	90.3	154	2	US-08-710-330A-11
4	696	87.7	154	3	US-08-817-787-18
5	584	73.6	113	2	US-08-710-330A-2
6	525	66.1	113	2	US-08-710-330A-1
7	515	64.9	113	2	US-08-710-330A-5
8	505	63.6	113	2	US-08-710-330A-4
9	504	63.5	113	2	US-08-710-330A-3
10	240	30.2	60	3	US-08-817-787-22
11	240	30.2	64	3	US-08-817-787-24
12	213	26.8	57	3	US-08-817-787-20
13	162	20.4	141	1	US-08-443-890-18
14	162	20.4	141	1	US-08-443-890-21
15	162	20.4	141	2	US-08-484-688B-65
16	162	20.4	141	3	US-08-463-160B-66
17	162	20.4	141	5	PCT-US92-0972-18
18	152	19.1	141	3	US-09-058-562-18
19	150	18.9	146	4	US-09-353-719-1
20	140	17.6	141	2	US-08-627-173-21
21	140	17.6	141	2	US-08-535-889A-21
22	140	17.6	141	3	US-08-424A-7
23	140	17.6	141	3	US-09-005-546-21
24	140	17.6	141	4	US-08-477-669-7
25	140	17.6	141	4	US-10-128-581-26
26	135	17.0	146	2	US-08-619-708A-4
27	16.1	141	2	US-08-627-173-19	

SEQUENCE

RESULT 1

US-08-932-176-7

; Sequence 7, Application US/08992176

; Patent No. 6125331

; GENERAL INFORMATION:

; APPLICANT: TOH, Hiroyuki

; TITLE OF INVENTION: STRUCTURAL ALIGNMENT METHOD MAKING USE OF A DOUBLE

; FILE REFERENCE: 9200-0001-2

CURRENT APPLICATION NUMBER: US/08/9932-176

CURRENT FILING DATE: 1997-12-17

EARLIER APPLICATION NUMBER: JP 8-340727

EARLIER FILING DATE: 1996-12-20

NUMBER OF SEQ ID NOS: 10

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO: 7

LENGTH: 153

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence:alpha-protein

OTHER INFORMATION: (lmbc)

US-08-992-176-7

Query Match 100.0% ; Score 794; DB 3;

Best Local Similarity 100.0% ; Pred. No. 2e-81;

Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSSEGEMQLVHLWVAKEADYAGHGDILRLFKSHPETLEKFDRPKHLTEAEMKASED 60

Db 1 VLSSEGEMQLVHLWVAKEADYAGHGDILRLFKSHPETLEKFDRPKHLTEAEMKASED 60

Qy 61 LKKHGTVLTAIGAILKKGGHEALKLPLAQSHATHKHPIKYLEFISEAIIHVLRHP 120

Db 61 LKKHGTVLTAIGAILKKGGHEALKLPLAQSHATHKHPIKYLEFISEAIIHVLRHP 120

RESULT 2

US-08-932-176-10

; Sequence 10, Application US/08992176

; Patent No. 6125331

; GENERAL INFORMATION:

; APPLICANT: TOH, Hiroyuki

; TITLE OF INVENTION: STRUCTURAL ALIGNMENT METHOD MAKING USE OF A DOUBLE

; FILE REFERENCE: 9200-0001-2

CURRENT APPLICATION NUMBER: US/08/992.176
 CURRENT FILING DATE: 1997-12-17
 EARLIER APPLICATION NUMBER: JP 8-340727
 EARLIER FILING DATE: 1996-12-20
 NUMBER OF SEQ ID NOS: 10
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO: 10
 LENGTH: 152
 TYPE: PFT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence:alpha-protein
 OTHER INFORMATION: (Imbc)
 US-08-992-176-10

Query Match 98.0%; Score 778.5; DB 3; Length 152;
 Best Local Similarity 99.3%; Pred. No. 1.e-79;
 Matches 152; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 VLSGEWOLVLYWAKEADVAHGQDQLIRLFKSHPTLEKFDKHLKTEAMKASED 60
 Db 1 VLSGEWOLVLYWAKEADVAHGQDQLIRLFKSHPTLEKFDKHLKTEAMKASED 60

QY 61 LKXHGTVLTALGAILKKKGHHEAELKPLAQSHATKHKPIKYLEFSEATHVLHSHP 120
 Db 61 LKXHGTVLTALGAILKKKGHHEAELKPLAQSHATKHKPIKYLEFSEATHVLHSHP 119

QY 121 GDFGADAQAMNKALELLFRKDIAKYKELGYQG 153
 Db 120 GDFGADAQAMNKALELLFRKDIAKYKELGYQG 152

RESULT 3
 US-08-710-330A-11
 Sequence 11, Application US/08710330A
 Patent No. 5854041

GENERAL INFORMATION:
 APPLICANT: Brayer, Gary D.
 APPLICANT: Lee, Hung
 APPLICANT: Mauk, Grant A.
 APPLICANT: Smith, Michael
 APPLICANT: Tong, Harry
 APPLICANT: Wan, Liangju
 TITLE OF INVENTION: MYOGLOBIN WITH PEROXIDASE ACTIVITY
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell & Flores LLP
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92122

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/710,330A
 FILING DATE: 16-SEP-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: GAY, David A.
 REGISTRATION NUMBER: 39,200
 REFERENCE/DOCKET NUMBER: P-SM 2262
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 154 amino acids
 TYPE: amino acid
 STRANDEDNESS:

Query Match 90.3%; Score 717; DB 2; Length 154;
 Best Local Similarity 88.2%; Pred. No. 9e-73;
 Matches 134; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

QY 2 LSGEWOVLHYWAKEADVAHGQDQLIRLFKSHPTLEKFDKHLKTEAMKASED 61
 Db 3 LSDGEWQVLNYWGVEADHGQBVJRLTGHPTELEKFDKHLKTEAMKASED 62

QY 62 KKHGTWTALGAILKKKGHHEAELKPLAQSHATKHKPIKYLEFISPAIIHVLSRHPG 121
 Db 63 KKHGTWTALGAILKKKGHHEAELKPLAQSHATKHKPIKYLEFISDAIIHVLSKHPG 122

QY 122 DFGADAQAMNKALELLFRKDIAKYKELGYQG 153
 Db 123 DFGADAQAMNKALELLFRKDIAKYKELGYQG 154

RESULT 4
 US-08-817-787-18
 Sequence 18, Application US/08817787
 Patent No. 6293353

GENERAL INFORMATION:
 APPLICANT: Iupas, Andrei
 APPLICANT: Pack, Peter
 TITLE OF INVENTION: TARGETED HETERO-ASSOCIATION OF RECOMBINANT PROTEINS TO MULTI-FUNCTIONAL COMPLEXES
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FISH & NEAVE
 STREET: 1251 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10200

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/817,787
 FILING DATE: 23-SEP-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP95/04117
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: EP 94 11 6558.1
 FILING DATE: 20-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Haley Jr., James F.
 REGISTRATION NUMBER: 27,794
 REFERENCE/DOCKET NUMBER: MORPHO/1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-596-9000
 TELEFAX: 212-596-9000
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 154 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Homo sapiens

US-08-817-787-18

Query Match 87.7%; Score 696; DB 3; Length 154;
 Best Local Similarity 84.2%; Pred. No. 2e-70;
 Matches 128; Conservative 14; Mismatches 10; Indels 0; Gaps 0;

RESULT 5		US-08-710-330A-2	
Qy	2 LSEGEWQLVLYHWAKEADVAGHQDILIRLFLKSHPETLEKFDRFKHLKTEAMKASEDL 61	;	Sequence 2, Application US/08710330A
;	Patent No. 5855041	;	
Db	3 LSDGEWQLVLYNWRGEADVAGHQDILIRLFLKSHPETLEKFDRFKHLKTEAMKASEDL 62	;	
Qy	62 KKHGTVTALGAIKKGGHEAEKLPLAQSHATKHKIPKYLEFISEAIIHVVLHSRHPG 1.21	;	
Db	63 KKHGTVTALGAIKKGGHEAEKLPLAQSHATKHKIPKYLEFISECIIQVLQSKHPG 1.22	;	
Qy	122 DFGADAOGMNNKAELFLRKDIAAKYKELGYCG 153	;	
Db	123 DFGADAEGMNNKAELFLRKDMASNYKELGYCG 154	;	
GENERAL INFORMATION:			
;	APPLICANT: Brayer, Gary D.	;	
;	APPLICANT: Lee, Hung	;	
;	APPLICANT: Mauk, Grant A.	;	
;	APPLICANT: Smith, Michael	;	
;	APPLICANT: Tong, Harry	;	
;	APPLICANT: Wan, Liangfu	;	
;	TITLE OF INVENTION: MYOGLOBIN WITH PEROXIDASE ACTIVITY	;	
;	NUMBER OF SEQUENCES: 11	;	
;	CORRESPONDENCE ADDRESS:	;	
;	ADDRESSEE: Campbell & Flores LLP	;	
;	STREET: 4370 La Jolla Village Drive, Suite 700	;	
;	CITY: San Diego	;	
;	STATE: California	;	
;	COUNTRY: USA	;	
;	ZIP: 92122	;	
;	COMPUTER READABLE FORM:	;	
;	MEDIUM TYPE: Floppy disk	;	
;	COMPUTER: IBM PC compatible	;	
;	OPERATING SYSTEM: PC-DOS/MS-DOS	;	
;	SOFTWARE: Patentin Release #1.0, Version #1.30	;	
;	CURRENT APPLICATION DATA:	;	
;	APPLICATION NUMBER: US/08/710,330A	;	
;	FILING DATE: 16-SEP-1996	;	
;	CLASSIFICATION: 435	;	
;	ATTORNEY/AGENT INFORMATION:	;	
;	NAME: GAY, David A.	;	
;	REGISTRATION NUMBER: 39,200	;	
;	REFERENCE/DECET NUMBER: P-SM 2262	;	
;	TELECOMMUNICATION INFORMATION:	;	
;	TELEPHONE: (619) 535-9001	;	
;	TELEFAX: (619) 535-8949	;	
;	INFORMATION FOR SEQ ID NO: 2:	;	
;	SEQUENCE CHARACTERISTICS:	;	
;	LENGTH: 113 amino acids	;	
;	TYPE: amino acid	;	
;	STRANDEDNESS:	;	
;	TOPOLOGY: linear	;	
;	US-08-710-330A-2	;	
;	Query Match 73.6%; Score 584; DB 2; Length 113;	;	
;	Best Local Similarity 100.0%; Pred. No. 5e-38; Mismatches 0; Indels 0; Gaps 0	;	
;	Matches 113; Conservative 0; Mis matches 0; Indels 0; Gaps 0	;	
Qy	1 VLSEGEGWQLVLYHWAKEADVAGHQDILIRLFLKSHPETLEKFDRFKHLKTEAMKASEDL 60	;	
Db	1 VLSEGEGWQLVLYNWRGEADVAGHQDILIRLFLKSHPETLEKFDRFKHLKTEAMKASEDL 60	;	
Qy	61 LKKHGTVTALGAIKKGGHEAEKLPLAQSHATKHKIPKYLEFISEAIIHVVLHSRHPG 1.13	;	
Db	61 LKKHGTVTALGAIKKGGHEAEKLPLAQSHATKHKIPKYLEFISEAIIHVVLHSRHPG 1.13	;	

COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/710,330A
 FILING DATE: 16-SEP-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: GAY, David A.
 REGISTRATION NUMBER: 39,200
 REFERENCE/DOCKET NUMBER: P-SM 2262
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 113 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-710-330A-5

Query Match 64.9%; Score 515; DB 2; Length 113;
 Best Local Similarity 85.6%; Pred. No. 2.8e-50;
 Matches 95; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Qy 2 LSBEWQLVHWAKVEDAVGHGQDILRLFKSHPTLEKDFRKHKTAEAKSDDL 61
 Db 2 LSBEWQLVHWAKVEDAVGHGQDILRLFKSHPTLEKDFRKHKTAEAKSDDL 61

Qy 62 KKHGATVLTALGALLKKGSHHEAKPLAQSHATKHKIPKYLEFISAIH 112
 Db 62 KKHGATVLTALGALLKKGSHHEAKPLAQSHATKHKIPKYLEFISAIH 112

RESULT 8
 US-08-710-330A-4
 ; Sequence 4, Application US/08710330A
 ; Patent No. 5854041
 GENERAL INFORMATION:
 APPLICANT: Brayer, Gary D.
 APPLICANT: Lee, Hung
 APPLICANT: Mauk, Grant A.
 APPLICANT: Smith, Michael
 APPLICANT: Tong, Harry
 APPLICANT: Wan, Liang
 TITLE OF INVENTION: MYOGLOBIN WITH PEROXIDASE ACTIVITY
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell & Flores LLP
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/710,330A
 FILING DATE: 16-SEP-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: GAY, David A.
 REGISTRATION NUMBER: 39,200
 REFERENCE/DOCKET NUMBER: P-SM 2262
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 113 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-710-330A-3

Query Match 63.6%; Score 505; DB 2; Length 113;
 Best Local Similarity 83.0%; Pred. No. 3.7e-49;
 Matches 93; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

Qy 2 LSBEWQLVHWAKVEDAVGHGQDILRLFKSHPTLEKDFRKHKTAEAKSDDL 61
 Db 2 LSBEWQLVHWAKVEDAVGHGQDILRLFKSHPTLEKDFRKHKTAEAKSDDL 61

Qy 62 KKHGATVLTALGALLKKGSHHEAKPLAQSHATKHKIPKYLEFISAIH 113
 Db 62 KKHGATVLTALGALLKKGSHHEAKPLAQSHATKHKIPKYLEFISAIH 113

RESULT 9
 US-08-710-330A-3
 ; Sequence 3, Application US/08710330A
 ; Patent No. 5854041
 GENERAL INFORMATION:
 APPLICANT: Brayer, Gary D.
 APPLICANT: Lee, Hung
 APPLICANT: Mauk, Grant A.
 APPLICANT: Smith, Michael
 APPLICANT: Tong, Harry
 APPLICANT: Wan, Liang
 TITLE OF INVENTION: MYOGLOBIN WITH PEROXIDASE ACTIVITY
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Campbell & Flores LLP
 STREET: 4370 La Jolla Village Drive, Suite 700
 CITY: San Diego
 STATE: California
 COUNTRY: USA
 ZIP: 92122
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/710,330A
 FILING DATE: 16-SEP-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: GAY, David A.
 REGISTRATION NUMBER: 39,200
 REFERENCE/DOCKET NUMBER: P-SM 2262
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 535-9001
 TELEFAX: (619) 535-8949
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 113 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-710-330A-3

Query Match 63.5%; Score 504; DB 2; Length 113;
 Best Local Similarity 83.9%; Pred. No. 4.8e-49;
 Matches 94; Conservative 9; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LSBEWQLVHWAKVEDAVGHGQDILRLFKSHPTLEKDFRKHKTAEAKSDDL 61

Db 2 LSDGEWQAVLNAGKVEADYAGHQEVLLRLFTGHPETLEKDFKPKHLKTEAMKASEDL 61
 Qy 62 KKHGVTVLTLGALLKKKSHEARPLKPLAQASHATKHKPIKYLEFISDAIILH 113
 Db 62 KKHGVTVLTLGALLKKKSHEARPLKPLAQASHATKHKPIKYLEFISDAIILH 113

RESULT 10
 US-08-817-787-22
 Sequence 22, Application US/08817787
 Patent No. 6294353
 GENERAL INFORMATION:
 APPLICANT: Pack, Peter
 APPLICANT: Lupas, Andrei
 TITLE OF INVENTION: TARGETED HETERO-ASSOCIATION OF RECOMBINANT PROTEINS TO MULTI-FUNCTIONAL COMPLEXES
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESS: FISH & NEAVE
 STREET: 1251 Avenue of the Americas
 CITY: New York
 STATE: New York
 ZIP: 10020
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/817,787
 FILING DATE: 23-SEP-1997
 CLASSIFICATION: 424
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: PCT/EP95/04117
 FILING DATE:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: EP 94 11 6558.1
 FILING DATE: 20-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Haley Jr., James F.
 REGISTRATION NUMBER: 27,794
 REFERENCE DOCKET NUMBER: MORPHO/1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-596-9000
 TELEFAX: 212-596-9090
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 64 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-817-787-24

Query Match 30.2%; Score 240; DB 3; Length 64;
 Best Local Similarity 80.0%; Pred. No. 9.4e-20;
 Matches 44; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 99 IPIKYRPISEAIHVILHSRHSRHFQDAGQAMMKALEFRKDIAKYKELGYQG 153
 Db 3 IPIKYRPISEAIHVILHSRHSRHFQDAGQAMMKALEFRKDIAKYKELGYQG 57

RESULT 12
 US-08-817-787-20
 Sequence 20, Application US/08817787
 Patent No. 6294353
 GENERAL INFORMATION:
 APPLICANT: Lupas, Andrei
 TITLE OF INVENTION: TARGETED HETERO-ASSOCIATION OF RECOMBINANT PROTEINS TO MULTI-FUNCTIONAL COMPLEXES
 NUMBER OF SEQUENCES: 36
 CORRESPONDENCE ADDRESS:
 ADDRESS: FISH & NEAVE
 STREET: 1251 Avenue of the Americas
 CITY: New York
 STATE: New York
 ZIP: 10020
 COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/817,787
 FILING DATE: 23-SEP-1997
 CLASSIFICATION: 424
 PRIORITY APPLICATION DATA:

RESULT 11
 US-08-817-787-24
 Sequence 24, Application US/08817787
 Patent No. 6294353
 GENERAL INFORMATION:
 APPLICANT: Pack, Peter
 APPLICANT: Lupas, Andrei
 TITLE OF INVENTION: TARGETED HETERO-ASSOCIATION OF RECOMBINANT PROTEINS TO MULTI-FUNCTIONAL COMPLEXES
 NUMBER OF SEQUENCES: 36

APPLICATION NUMBER: PCT/EP95/04117
 FILING DATE:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: EP 94 11 6558.1
 FILING DATE: 20-OCT-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Haley Jr., James F.
 REGISTRATION NUMBER: 27,794
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-596-9000
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 57 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-817-787-20

Query	Match	Score	Length
	Query Match	26.8%	DB 3;
	Best Local Similarity	80.4%	Fred. No. B.6e-17;
	Matches 37;	Conservative 6;	Mismatches 3;
			Indels
Qy	2 LSEGEWQLVLIHWKVEADVAGHGDIDLIRIFKSHPTLEKFDKFK 47		
Ddb	5 LSDGEWQLVLIWGRYEADIFGHQGEVLRLIRIFKSHPTLEKFDKFK 50		

RESULT 13
 US-08-240-712-18
 Sequence 18, Application US/08240712
 Patient No. 5599907
 GENERAL INFORMATION:
 APPLICANT: ANDERSON, DAVID C.
 APPLICANT: MATHERS, ANTHONY JAMES
 APPLICANT: STETLER, GARY L.
 TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
 TITLE OF INVENTION: HEMOGLOBINS
 NUMBER OF SEQUENCES: 35
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Browdy and Neimark
 STREET: 419 Seventh Street, N.W., Suite 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/240,712
 FILING DATE: 09-MAY-1994
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/09752
 FILING DATE: 13-MAY-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: COOPER, TIVER P
 REGISTRATION NUMBER: 28,005
 TELECOMMUNICATION INFORMATION:
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEEN: single
 TOPOLOGY: linear

US-08-240-712-18 MOLECULE TYPE: peptide

Query Match	20
Best Local Similarity	28
Matches 41; Conservative	
Qy	2 LSGEWQVLVHWA : : :
Db	2 LTKTERTIVSMWAK
Qy	62 KKHGVTVLTA : : : :
Db	58 -HSSKVVAVGDAV
Qy	122 DFGDAGGMNKALE : : :
Db	116 DFTPEAHAAWDKFU

RESULT 14
 US-08-443-890-18
 ; Sequence 18, Application US
 ; Patent No. 5739011
 ; GENERAL INFORMATION:
 ; APPLICANT: ANDERSON, DE
 ; ADDRESS: 1000 N. 100 E.
 ; CITY: Salt Lake City
 ; STATE: UT
 ; COUNTRY: USA
 ; ZIP: 84111
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY DISK
 ; COMPUTER: IBM PC compat
 ; OPERATING SYSTEM: PC
 ; SOFTWARE: Patent In Re
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US
 ; FILING DATE: 31-MAY-1997
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US
 ; FILING DATE: 09-MAY-1997
 ; APPLICATION NUMBER: US
 ; FILING DATE: 13-MAY-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: COOPER, TIVER P
 ; REGISTRATION NUMBER:
 ; REFERENCE/DOCKET NUMBER:
 ; TELECOMMUNICATIONS INFORMATION:
 ; TELEPHONE: 202-628-5132
 ; TELEX: 228-737-3524
 ; INFORMATION FOR SEQ ID NO
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 141 amino acid
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 US-08-443-890-18
 Query Match 20
 Best Local Similarity 28
 Matches 41; Conservative

RESULT 15
US-08-484-686B-65

Sequence 65, Application US/08484686B
Patent No. 5827693

GENERAL INFORMATION:

APPLICANT: De Angelis, Joseph
APPLICANT: Motwani, Nalini
APPLICANT: Bajwa, Waheed

TITLE OF INVENTION: Expression of Recombinant Hemoglobin and
NUMBER OF SEQUENCES: 71

CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,686B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/368,407
FILING DATE: 29-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/876,290
FILING DATE: 29-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/684,611
FILING DATE: 12-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Plincoff, Gerald J.
REGISTRATION NUMBER: 20,823
REFERENCE/DOCKET NUMBER: 6666-043-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 869-8864
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-484-686B-65

Query Match 20.4%; Score 162; DB 2; Length 141;
Best Local Similarity 28.1%; Pred. No. 1.6e-10;
Matches 41; Conservative 29; Mismatches 6; Gaps 1;

Y 2 LSEGEWOLVHWAKEYADVAHGQDILIRLFKSHPETLEKFDRFKHLKTEAMKASEDL 61
Db 2 LTKTERTIVSMWAKISTQADTTGTCETEFLPFLSHPQTKYFPHEDLHPGSQQLRA --- 57

Y 62 KKHGVTVLTAALKKKKGHEAELKPLAQSHATKHKIPIKYLEFISEAIIHVLSRHFG 121
Db 58 --HGSKVVAVGDAVKSIDDIGGALKSUSELHAYILRVDPVNFKLSHCLLVTLAARPPA 115

Y 122 DFGADAOQAMNKALELFERDIARYK 147
Db 116 DFTAEAHAWDKFLSVSSVSLTEKRY 141

Search completed: August 10, 2004, 15:30:20
Job time : 8.77966 secs

B
D
N
T

Copyright GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 10, 2004, 15:28:46 ; Search time 22.2276 Seconds

Perfect score: 794

Sequence: VLSSEGEMQLVHLVWAKVEAD.....ALELFRKDIAAKYKELGYQG 153

Scoring table: BLOSUM62

Title: US-09-455-978B-76

Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing First 45 summaries

Database : Published Applications AA:
 1: /cgn2_6/ptodata/2/pubpaal/us07_PUBCOMB.pep:
 2: /cgn2_6/ptodata/2/pubpaal/PCT_NEW_PUB.pep:
 3: /cgn2_6/ptodata/2/pubpaal/us06_NEW_PUB.pep:
 4: /cgn2_6/ptodata/2/pubpaal/us05_PUBCOMB.pep:
 5: /cgn2_6/ptodata/2/pubpaal/us07_NEW_PUB.pep:
 6: /cgn2_6/ptodata/2/pubpaal/PCTUS_PUBCOMB.pep:
 7: /cgn2_6/ptodata/2/pubpaal/us08_NEW_PUB.pep:
 8: /cgn2_6/ptodata/2/pubpaal/us09_PUBCOMB.pep:
 9: /cgn2_6/ptodata/2/pubpaal/us09A_PUBCOMB.pep:
 10: /cgn2_6/ptodata/2/pubpaal/us09B_PUBCOMB.pep:
 11: /cgn2_6/ptodata/2/pubpaal/us09C_PUBCOMB.pep:
 12: /cgn2_6/ptodata/2/pubpaal/us09_NEW_PUB.pep:
 13: /cgn2_6/ptodata/2/pubpaal/us10_PUBCOMB.pep:
 14: /cgn2_6/ptodata/2/pubpaal/us10B_PUBCOMB.pep:
 15: /cgn2_6/ptodata/2/pubpaal/us10C_PUBCOMB.pep:
 16: /cgn2_6/ptodata/2/pubpaal/us10_NEW_PUB.pep:
 17: /cgn2_6/ptodata/2/pubpaal/us60_NEW_PUB.pep:
 18: /cgn2_6/ptodata/2/pubpaal/us60_PUBCOMB.pep:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

%

Result No.	Score	Query	Match Length	DB ID	Description
1	696	87.7	154	16 US-10-408-765A-106	Sequence 106, App
2	162	20.4	141	9 US-09-977-577-19	Sequence 19, Appl
3	152	19.1	31	15 US-10-289-009-17	Sequence 17, Appl
4	140	17.6	141	9 US-09-839-164-5	Sequence 7, Appl
5	140	17.6	141	14 US-10-128-51-26	Sequence 26, Appl
6	133	16.8	141	16 US-10-463-59-28	Sequence 28, Appl
7	133	16.8	31	15 US-10-289-009-4	Sequence 4, Appl
8	133	16.8	154	15 US-10-378-019-102	Sequence 102, Appl
9	132	16.6	142	12 US-10-424-599-22747	Sequence 22747,
10	132	16.6	142	14 US-10-205-219-189	Sequence 189, Appl
11	128	16.1	141	9 US-09-839-164-5	Sequence 5, Appl
12	128	16.1	141	14 US-10-128-581-24	Sequence 24, Appl
13	128	16.1	141	16 US-10-463-699-26	Sequence 26, Appl
14	127	16.0	146	9 US-09-839-164-8	Sequence 8, Appl
15	127	16.0	146	14 US-10-128-581-27	Sequence 27, Appl

SEQUENCES

16 US-10-463-699-29

17 125 15.7 146 9 US-09-977-577-20

18 124 15.6 146 9 US-09-977-577-17

19 120 15.1 146 9 US-09-839-164-6

20 120 15.1 146 14 US-10-128-581-25

21 120 15.1 146 16 US-10-463-699-27

22 114 14.4 147 9 US-09-117-490-8

23 112 14.1 141 9 US-09-839-164-2

24 112 14.1 141 9 US-09-977-577-14

25 112 14.1 141 9 US-09-977-577-18

26 112 14.1 141 12 US-10-280-725B-8

27 112 14.1 141 13 US-10-085-853-31

28 112 14.1 141 14 US-10-128-561-21

29 112 14.1 141 14 US-10-280-679B-9

30 112 14.1 141 16 US-10-463-699-23

31 112 14.1 142 15 US-10-435-666-7

32 111 14.0 141 16 US-10-408-765A-310

33 104 13.1 122 15 US-10-115-482-52

34 103 13.0 146 9 US-09-839-164-4

35 103 13.0 146 9 US-09-977-577-15

36 103 13.0 146 13 US-10-085-853-33

37 103 13.0 146 14 US-10-128-591-23

38 103 13.0 146 16 US-10-463-699-25

39 103 13.0 146 16 US-10-408-765A-3019

40 101 12.7 19 15 US-10-289-009-12

41 101 12.7 20 12 US-10-103-395-189

42 100 12.6 147 9 US-09-977-577-21

43 99 12.5 146 9 US-09-977-577-16

44 98 12.3 146 16 US-10-408-765A-84

45 96 12.1 146 15 US-10-289-009-13

ALIGMENTS

RESULT 1 US-10-408-765A-106

; Publication No. US10408765A1

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Boin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME

; CURRENT APPLICATION NUMBER: US/10/408,765A

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO: 106

; LENGTH: 154

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-10-408-765A-106

Query Match Score 696; DB 16; Length 154;

Best Local Similarity 84.2%; Pred. No. 2.3e-69;

Matches 128; Conservative 14; Mismatches 10; Indels 0; Gaps 0;

Qy 2 LSEGEVQLVTHWAKVEADYAGGQDILRFLKSHPTETLKFDKFLKTEAEMKASEDL 61

Db 3 LSDGEWQLVLNWGKTEAD PGHGOEVLRLFKGHPTELKFDKFLKSDEMKASEDL 62

Qy 62 KKHGVTVLTAIGAIKKKGHAEALKPLAQSHATKHKIPKYLIFLSELAIIHVHLRSRPG 121

Db 63 KRHGAVLTAIGGIIRKGHAEALKPLAQSHATKHKIPKYLEFTSCTIVLQSCHKPG 122

Qy 122 DFGADAGQAMNKALELFRKDIAAKYKELGYQG 153

Db 123 DFGADAGGMKALELFRKDMASNYKEIGFQG 154

RESULT 2
US-09-977-577-19

Sequence 19, Application US/09977577

GENERAL INFORMATION:

APPLICANT: MOESTRUE, Soren

PATENT NO.: US20020155995A1

PRIOR FILING DATE: 2001-02-22

PRIOR APPLICATION NUMBER: DK PA 2001 00039

PRIOR FILING DATE: 2001-01-11

PRIOR APPLICATION NUMBER: DK PA 2000 01543

PRIOR FILING DATE: 2000-10-16

NUMBER OF SEQ ID NOS: 25

SEQ ID NO: 19

LENGTH: 141

TYPE: PRT

ORGANISM: HOMO sapiens

US-09-977-577-19

Query Match 20.4%; Score 162; DB 9; Length 141;

Best Local Similarity 28.1%; Pred. No. 7.4e-10;

Matches 41; Conservative 29; Mismatches 70; Indels 6; Gaps 1;

Qy 2 LSGEWEQVILHWAKEADVAHGQDILIRLFKSHPETLEKDFRKHKTEAMKASEDL 61

Db 2 LTKTERTIVSWMKISTOATDTIGTERERFLPFLSHPQTKYTFPHFDLHPGSQQLRA --- 57

Qy 62 KKHGVTVLTAITGAILKKKGHEAELPKLAQSHATKHKPIKYLEFISIAITHVLHSRPG 121

Db 58 --HGSKTTVAVGDAVKSDDIGGALKSKSELHAYILRVDPVNFKLSCLLVTAAFPKA 115

Qy 122 DFGADAGGMKALELFRKDIAKYK 147

Db 116 DFTAEAHAWDKPLSVWSVSLTEKYR 141

RESULT 3
US-10-289-009-17

Sequence 17, Application US/10289009

GENERAL INFORMATION:

APPLICANT: Peters, Eric C.

APPLICANT: Brock, Ansgar

APPLICANT: Ericson, Christer

APPLICANT: IRM LLC

TITLE OF INVENTION: Labeling Reagent and Methods of Use

FILE REFERENCE: 021288-000230US

CURRENT APPLICATION NUMBER: US/10/289,009

CURRENT FILING DATE: 2003-04-01

PRIOR APPLICATION NUMBER: US 60/332,989

PRIOR FILING DATE: 2001-11-05

PRIOR APPLICATION NUMBER: US 60/385,835

PRIOR FILING DATE: 2002-06-03

PRIOR APPLICATION NUMBER: US 60/410,382

PRIOR FILING DATE: 2002-09-12

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO: 17

LENGTH: 31

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: equine

OTHER INFORMATION: myoglobin tryptic polypeptide #16
US-10-289-009-17

Query Match 19.1%; Score 152; DB 15; Length 31;

Best Local Similarity 90.3%; Pred. No. 1.2e-09;

Matches 28; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 103 YLEFISDAITHVILHSRPGDFGADAGGMNK 133

Db 1 YLEFISDAITHVILHSRPGDFGADAGGMNK 31

RESULT 4
US-09-839-164-7

Sequence 7, Application US/09839164

PATENT NO.: US200200985B1A1

GENERAL INFORMATION:

APPLICANT: KOZLOV, VLADIMIR

TSYRLOVA, IRENA

TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND THE USES THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYE P. C.

STREET: 1100 NORTH GLEBE ROAD, 8th FLOOR

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: U.S.A.

ZIP: 22201-4714

COMPUTER READABLE FORM:

MEDIUM TYPE: 1.44 Mb diskette

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/Ms-DOS

SOFTWARE: MS Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/839,164

FILING DATE: 3-APR-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/477,668

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: US 08/316,424

FILING DATE: 30-SEP-1994

APPLICATION NUMBER: PCT/US94/03349

FILING DATE: 29-MAR-1994

APPLICATION NUMBER: US 08/040,942

FILING DATE: 31-MAR-1993

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 141 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-839-164-7

Query Match 17.6%; Score 140; DB 9; Length 141;

Best Local Similarity 27.6%; Pred. No. 2.1e-07;

Matches 42; Conservative 20; Mismatches 74; Indels 16; Gaps 4;

Qy 1 VLSGEWEQVILHWAKEADVAHGQDILIRLFKSHPETLEKDFR -KHLKTEAEMKAS 58

Db 1 VLSAADKANVKAAWGVQAGHGAELERNFLGFPTKTYFPHFNLSH----GS 52

Qy 59 EDLURGHGVTVLTAITGAILKKKGHEAELPKLAQSHATKHKPIKYLEFISIAITHV 115

Db 53 DQVKAHGGQKVADAL--TKAVGHLDLPGALSALSDLHAKLVRDPYNFKLSCLLVLT 109

Qy 116 HSHPGDFGADAGGMNKALFRKDIAKYK 147

Db 110 AAHPDDNNPSVHASLDKPLANVSTVLSKYR 141

RESULT 5
 US-10-128-581-26
 Sequence 26, Application US/10128581
 GENERAL INFORMATION:
 ADDRESSEE: TSYRLOVA, IRENA
 APPLICANT: WOLPE, STEPHEN D.
 TITLE OF INVENTION: HEMOGLOBIN ALPHA CHAIN PEPTIDE FRAGMENTS USEFUL FOR INHIBITING STEM CELL PROLIFERATION

NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 1.44 MB Disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS Word

CURRENT APPLICATION DATA:
 FILING DATE: 24-Apr-2002
 CLASSIFICATION: <Unknown>
 PRIORITY APPLICATION DATA:
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/128,581
 FILING DATE: 24-Apr-2002
 CLASSIFICATION: <Unknown>
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 09/617,840
 FILING DATE: 17-JUL-2000
 APPLICATION NUMBER: US 09/005,546
 FILING DATE: 12-JAN-1998
 APPLICATION NUMBER: US 08/535,882
 FILING DATE: 28-SEP-1995
 APPLICATION NUMBER: US 08/316,424
 FILING DATE: 30-SEP-1994
 APPLICATION NUMBER: PCT/US94/03349
 FILING DATE: 29-MAR-1994
 APPLICATION NUMBER: US 08/040,942
 FILING DATE: 31-MAR-1993
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 28:
 US-10-463-699-28

Query Match Score 140; DB 16; Length 141;
 Best Local Similarity 27.6%; Pred. No. 2.1e-07;
 Matches 42; Conservative 20; Mismatches 74; Indels 16; Gaps 4;

QY 1 VLSEGEWOLVLHWAKEADYAGHGQDILIRLFKSHPETLEKFDRF-KHLKTEAEMKAS 58
 Db 1 VLSAAAKVKAAGVGGQAGAHGEALERMFGLPPTTYFPHFLNSH-----GS 52

QY 59 EDLKKHGTVTALGALKKGHH--AEIPLPLASHATAFKHPIKYLLEFISEALIHVL 115
 Db 53 DQVKANGSQKVADL--TKAVGHLDDLGPGALSALSDHLAKLRVDPVNFKLISHCLVTL 109

Qy 116 HSRHPGDFGADAQGMNKALEBLFRKDAAKYK 147
 Db 110 AAHHPPDDFNPSVHASLDKFLANVSTVLTSKYR 141

RESULT 7
 US-10-289-009-4
 Sequence 4, Application US/10289009
 GENERAL INFORMATION:
 Publication No. US2003022870A1
 ;
 ; APPLICANT: Peters, Eric C.
 ; APPLICANT: Brock, Ansgar
 ; APPLICANT: Erickson, Christer
 ; APPLICANT: IRM LLC
 ; TITLE OF INVENTION: Labeling Reagent and Methods of Use
 ; FILE REFERENCE: 021288-000230US

RESULT 6
 US-10-463-699-28
 Sequence 28, Application US/10463699
 GENERAL INFORMATION:
 Publication No. US20040081640A1
 ;
 ; APPLICANT: TSYRLOVA, IRENA

CURRENT APPLICATION NUMBER: US/10/289,009
 CURRENT FILING DATE: 2003-04-01
 PRIOR APPLICATION NUMBER: US 60/332,988
 PRIOR FILING DATE: 2001-11-05
 PRIOR APPLICATION NUMBER: US 60/385,835
 PRIOR FILING DATE: 2002-06-03
 PRIOR APPLICATION NUMBER: US 60/410,382
 PRIOR FILING DATE: 2002-09-12
 NUMBER OF SEQ ID NOS: 29
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 4
 LENGTH: 31
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE: Description of Artificial Sequence:equine
 OTHER INFORMATION: myoglobin tryptic polypeptide #3
 us-10-289-009-4

Query Match
 Best Local Similarity 76.7%; Pred. No. 1.6e-07;
 Matches 23; Conservative 5; MisMatches 2; Indels 0; Gaps 0;
 RESULT 8
 US-10-378-029-102
 Sequence 102, Application US/10378029
 PUBLICATION NUMBER: US20040014087A1
 APPLICANT: HODGSON, David M.; LINCOLN, Stephen E.
 APPLICANT: RUSSO, Frank D.; SPIRO, Peter A.
 APPLICANT: BANVILLE, Steve C.; BRATCHER, Shawn R.
 APPLICANT: DOFOUR, Gerard B.; COHEN, Howard J.
 APPLICANT: ROSEN, Bruce; CHALUP, Michael S.
 APPLICANT: JACKSON, Jennifer L.; JONES, Annisa L.
 APPLICANT: YU, Jimmy Y.; GREENAWALT, Lila B.
 APPLICANT: PANZER, Scott R.; ROSEBERRY, LINCOLN, Ann M.
 APPLICANT: WRIGHT, Rachel J.; DANIELS, Susan E.
 TITLE OF INVENTION: MOLECULES FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: PT-1022-1.CIP
 CURRENT FILING DATE: 2003-02-28
 PRIOR FILING DATE: 2001-11-30
 PRIOR APPLICATION NUMBER: PCT/US00/15404
 PRIOR FILING DATE: 2000-05-31
 PRIOR APPLICATION NUMBER: US 60/147,500
 PRIOR FILING DATE: 1999-08-05
 PRIOR APPLICATION NUMBER: US 60/147,542
 PRIOR FILING DATE: 1999-08-05
 PRIOR APPLICATION NUMBER: US 60/147,541
 PRIOR FILING DATE: 1999-08-05
 PRIOR APPLICATION NUMBER: US 60/147,824
 PRIOR FILING DATE: 1999-08-05
 PRIOR APPLICATION NUMBER: US 60/147,547
 PRIOR FILING DATE: 1999-08-05
 PRIOR APPLICATION NUMBER: US 60/147,530
 PRIOR FILING DATE: 1999-08-05
 PRIOR FILING DATE: 1999-08-05
 PRIOR FILING DATE: 1999-08-05
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 104
 SOFTWARE: PERL Program
 SEQ ID NO 102
 LENGTH: 154
 TYPE: PRT
 ORGANISM: Homo sapiens

FEATURE:
 NAME/KEY: misc_feature
 OTHER INFORMATION: Incyte ID NO. US20040014087A1 441779_1.j.orfl
 US-10-378-029-102
 Query Match
 Best Local Similarity 26.7%; Pred. No. 1.4e-06;
 Matches 40; Conservative 24; MisMatches 74; Indels 12; Gaps 3;
 Qy 1 VLSGEWQVLVHWAKEYDVAHGQDILIRPKSHPTELEKPDFRFLKTEAEMKASED 60
 Db 14 VLSADDKTNKNQNCWKGKIGGGYGEALQRMFAAFPPTTKTYFISHIDVSPGSAQVKA-- 70
 Qy 61 LKKHGTVLTLGALKKKGHH--AEKLPLAQSHATKHKPIKYLEFISAIHVVLHS 117
 Db 71 --HGKXVADALA--KAADHIVEDLPGALSTSDLHAKLRVDPPNFKFLSHCLVTLAC 124
 Qy 118 RHFGDFGADAQGAMMKALELFRKDIARYK 147
 Db 125 HHPGDFTPAMHASLDKFSLASVSTVLTSKYR 154
 RESULT 9
 US-10-424-599-227247
 Sequence 227247, Application US/10424599
 PUBLICATION NUMBER: US2004031072A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa Thomas J
 APPLICANT: Kovacic David K
 APPLICANT: Zhou Yihua
 APPLICANT: Cao Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(5323)B
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 285684
 SEQ ID NO 227247
 LENGTH: 142
 TYPE: PRT
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_47233C.1.pep
 US-10-424-599-227247
 Query Match
 Best Local Similarity 26.7%; Pred. No. 1.7e-06;
 Matches 40; Conservative 24; MisMatches 74; Indels 12; Gaps 3;
 Qy 1 VLSGEWQVLVHWAKEYDVAHGQDILIRPKSHPTELEKPDFRFLKTEAEMKASED 60
 Db 2 VLSADDKTNKNQNCWKGKIGGGYGEALQRMFAAFPPTTKTYFISHIDVSPGSAQVKA-- 58
 Qy 61 LKKHGTVLTLGALKKKGHH--AEKLPLAQSHATKHKPIKYLEFISAIHVVLHS 117
 Db 59 --HGKXVADALA--KAADHIVEDLPGALSTSDLHAKLRVDPPNFKFLSHCLVTLAC 112
 Qy 118 RHFGDFGADAQGAMMKALELFRKDIARYK 147
 Db 113 HHPGDFTPAMHASLDKFSLASVSTVLTSKYR 142
 RESULT 10
 US-10-205-219-189
 Sequence 189, Application US/10205219
 Publication No. US2003013803A1
 GENERAL INFORMATION:
 APPLICANT: Warner-Lambert Company
 APPLICANT: Lee, Kevin
 APPLICANT: Dixon, Alishaar
 APPLICANT: Bicknabank, Robert
 APPLICANT: Pinnoch, Robert
 TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain

FILE REFERENCE: WL-A-018200
 CURRENT FILING DATE: US/10/205,219
 PRIOR APPLICATION NUMBER: GB 002-07-24
 PRIOR FILING DATE: 2002-07-27
 NUMBER OF SEQ ID NOS: 197
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 189
 LENGTH: 142
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 FEATURE:
 OTHER INFORMATION: Globin, alpha, major
 US-10-205-219-189

Query Match Score 16.1%; Best Local Similarity 27.0%; Pred. No. 4.6e-06; Matches 22; Mismatches 73; Indels 16; Gaps 4;
 Query Match Score 16.1%; Best Local Similarity 27.0%; Pred. No. 4.6e-06; Matches 41; Mismatches 73; Indels 16; Gaps 4;

Qy 1 VLSGEQVQLVWAKVEADYAGHGDILIRLFKSHPETLKFDRL - KHLXTEAMKAS 58
 Db 1 VLSGEQVLSNIKAAGKIGGHAEGGAELERMFASSPTTKTYFPHDVSH ---GS 52
 Qy 59 EDLKRHGKVTVTALGATLKCKKGHE --AEILPLAQSHATKHKIPKYLEFSEATHV 115
 Db 53 AQVKHQGKVKVADALAS --ARGHLDLPGALSALSULHANKLRVDVNFKLJSHCLVTL 109

Query Match Score 13.2; DB 14; Length 142;
 Best Local Similarity 26.7%; Pred. No. 1.7e-06;
 Matches 40; Conservative 24; Mismatches 74; Indels 12; Gaps 3;
 Qy 1 VLSGEQVQLVWAKVEADYAGHGDILIRLFKSHPETLKFDRLKHLXTEAMKASED 60
 Db 2 VLSADDKTNKNCWKGKIGGGEYGEALQRMFAAAPTPTKTYFSHIDVSQSAQYKA -- 58

RESULT 12
 US-10-128-581-24
 Sequence 24, Application US/10128581
 Publication No. US20030104984A1
 GENERAL INFORMATION:
 APPLICANT: TSYROLOVA, IRENA
 WOLPP, STEPHEN D.
 TITLE OF INVENTION: HEMOGLOBIN ALPHA CHAIN PEPTIDE FRAGMENTS USEFUL FOR INHIBITING STEM CELL PROLIFERATION
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHEY P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/128,581
 FILING DATE: 24-Apr-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/617,840
 FILING DATE: 17-JUL-2000
 APPLICATION NUMBER: US 09/005,546
 FILING DATE: 12-JAN-1998
 APPLICATION NUMBER: US 08/535,882
 FILING DATE: 28-SEP-1995
 APPLICATION NUMBER: US 08/316,424
 FILING DATE: 30-SEP-1994
 APPLICATION NUMBER: PCT/US94/03349
 FILING DATE: 29-MAR-1994
 APPLICATION NUMBER: US 08/040,942
 FILING DATE: 31-MAR-1993
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 24:
 US-10-128-581-24

Query Match Score 16.1%; Best Local Similarity 27.0%; Pred. No. 4.6e-06; Matches 22; Mismatches 73; Indels 16; Gaps 4;
 Query Match Score 16.1%; Best Local Similarity 27.0%; Pred. No. 4.6e-06; Matches 41; Conservative 22; Mismatches 73; Indels 16; Gaps 4;

FILE REFERENCE: WL-A-018200
 CURRENT FILING DATE: US/10/205,219
 PRIOR APPLICATION NUMBER: US 08/477,668
 FILING DATE: 07-JUN-1995
 APPLICATION NUMBER: US 08/316,424
 FILING DATE: 30-SEP-1994
 APPLICATION NUMBER: PCT/US94/03349
 FILING DATE: 29-MAR-1994
 APPLICATION NUMBER: US 08/040,942
 FILING DATE: 31-MAR-1993
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 141 amino acids
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear

Qy 1 VLSGEQWQLVHWWAKVEADYAGHGQDILIRLFKSHPETLEKPKDRF - KHLKTEAEMKAS 58
Db 1 VLSGEDKSNIKAANGKIGGGAAEYGAPELMFASTPTKTYFPHFDVSH-----GS 52

Qy 59 EDLJKHGTVTALGATLKKKGHE -- AELKPLAQSHATHKHPKYLEFISEAIIHVL 115
Db 53 AOVKGHRKVKVADLAS - - AELKPLAQSHATHKHPKYLEFISEAIIHVL 109

Qy 116 HSRHPGDGFADAQGMNKKALELFRKDIAKYK 147
Db 110 ASHPADFTPAVHASLDKFLASVSTVLTSKYR 141

RESULT 13
US-10-463-699-26
Sequence 26, Application US/10463699
Publication No. US2004008164041
GENERAL INFORMATION:
APPLICANT: WOLPE, STEPHEN D.
TITLE OF INVENTION: HEMOGLOBIN ALPHA CHAIN PEPTIDE FRAGMENTS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P. C.
STREET: 1100 NORTH GLEBEE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 MB Disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/463,699
FILING DATE: 18-JUN-2003
CLASSIFICATION: <Unknown>
PROR APPLICATION DATA:
APPLICATION NUMBER: US 10/128,581
FILING DATE: 24-APR-2002
APPLICATION NUMBER: US 09/617,840
FILING DATE: 17-JUL-2000
APPLICATION NUMBER: US 09/005,546
FILING DATE: 12-JAN-1998
APPLICATION NUMBER: US 08/535,882
FILING DATE: 28-SEP-1995
APPLICATION NUMBER: US 08/316,424
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: PCT/US94/03349
FILING DATE: 29-MAR-1994
APPLICATION NUMBER: PCT/US94/03349
FILING DATE: 31-MAR-1993
APPLICATION NUMBER: US 08/040,942
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 141 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-10-463-699-26

Query Match 16.1%; Score 128; DB 16; Length 141;
Best Local Similarity 27.0%; Pred. No. 4 6e-06; Mismatches 73; Indels 4;
Matches 41;

Qy 1 VLSGEQWQLVHWWAKVEADYAGHGQDILIRLFKSHPETLEKPKDRF - KHLKTEAEMKAS 58
Db 1 VLSGEDKSNIKAANGKIGGGAAEYGAPELMFASTPTKTYFPHFDVSH-----GS 52

Qy 59 EDLJKHGTVTALGATLKKKGHE -- AELKPLAQSHATHKHPKYLEFISEAIIHVL 115
Db 53 AOVKGHRKVKVADLAS - - AELKPLAQSHATHKHPKYLEFISEAIIHVL 109

Qy 116 HSRHPGDGFADAQGMNKKALELFRKDIAKYK 147
Db 110 ASHPADFTPAVHASLDKFLASVSTVLTSKYR 141

RESULT 14
US-09-839-164-8
Sequence 8, Application US/09839164
Patent No. US2002098583A1
GENERAL INFORMATION:
APPLICANT: KOZLOV, VLADIMIR
TSYRLOVA, IRENA
TITLE OF INVENTION: INHIBITOR OF STEM CELL PROLIFERATION AND
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P. C.
STREET: 1100 NORTH GLEBEE ROAD, 8th FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 Mb diskette
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS Word

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,164
FILING DATE: 23-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,668
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/316,424
FILING DATE: 30-SEP-1994
APPLICATION NUMBER: PCT/US94/03349
FILING DATE: 29-MAR-1994
APPLICATION NUMBER: US 08/040,942
FILING DATE: 31-MAR-1993
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-839-164-8

Query Match 16.0%; Score 127; DB 9; Length 146;
Best Local Similarity 27.6%; Pred. No. 6.2e-06; Mismatches 20; Indels 2; Gaps 1;
Matches 40;

Qy 2 LSEBQEWQLVHWWAKVEADYAGHGQDILIRLFKSHPETLEKPKDRFKHLPTEAEMKASEDL 61
Db 3 LSABKEAVLGLWKGKVNDV - GEALGRLIIVVYPWTQRFPFFSGDLSNADVMGNFKV 60

Qy 62 KKHGTVTALGATLKKKGHEAELKPLAQSHATHKHPKYLEFISEAIIHVLHSRHPG 121
Db 61 XANGKKVLOQSDFGKHDNLKGTPWTAELKPLAQSHATHKHPKYLEFISEAIIHVLARRLGH 120

Query Match 16.0%; Score 128; DB 9; Length 141;
Best Local Similarity 27.0%; Pred. No. 4 6e-06; Mismatches 73; Indels 4;
Matches 41;

Qy 122 DFGDAQGMNKKALELFRKDIAKY 146
Db 121 DFNEDVQAAFQKTVVAGVANALAHY 145

RESULT 15
US-10-128-581-27

Sequence 27, Application US/10128581
 GENERAL INFORMATION.
 APPLICANT: TSYRLOVA, IRENA
 TITLE OF INVENTION: HEMOGLOBIN ALPHA CHAIN PEPTIDE FRAGMENTS USEFUL
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: NIXON & VANDERHYE P.C.
 STREET: 1100 NORTH GLEBE ROAD
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MS WORD
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/128,581
 FILING DATE: 24-APR-2002
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 09/617,840
 FILING DATE: 17-JUL-2000
 APPLICATION NUMBER: US 09/005,546
 FILING DATE: 12-JAN-1998
 APPLICATION NUMBER: US 08/535,882
 FILING DATE: 28-SEP-1995
 APPLICATION NUMBER: US 08/316,424
 FILING DATE: 30-SEP-1994
 APPLICATION NUMBER: PCT/US94/03349
 FILING DATE: 29-MAR-1994
 APPLICATION NUMBER: US 08/040,942
 FILING DATE: 31-MAR-1993
 INFORMATION FOR SEQ ID NO: 27:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 146 amino acids
 TYPE: amino acid
 STRANDBNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 SEQUENCE DESCRIPTION: SEQ ID NO: 27:
 US-10-128-581-27

Query Match Score 127; DB 14; Length 146;
 Best Local Similarity 27.6%; Pred. No. 6 2e-06;
 Matches 40; Conservative 20; Mismatches 83; Indels 2; Gaps 1;
 Query 2 LSEGEGVQLVHWWAKYEADYAGHGQDILITLRFKSHIPETLEKFDRPKHLKEAEMKASEDI 61
 3 LSAEEKEAVIGLWGRKVNDEV- -GGPAGLGRILVVPPWTQRFESFGDLSNADAAGNPKV 60
 Query 62 KKHGIVTVLTALGAIKKKGHEAELKPLAQSHATKHKIPKYLEFISEALIHVLHSRPG 121
 Database 61 KAHRGKRVLQESDGKXKHLDDNKGTFAKLSLHCDQHVDENFRUJGNVIVVLLARRIGH 120
 122 DFGADAGQGANAKALEFRKDIAKY 146
 121 DFNPDVQAAFOKVVAGVANALAHY 145

Search completed: August 10, 2004, 15:42:27
 Job time : 23.2276 secs

B
I
A
N
K

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: August 10, 2004, 15:26:21 ; Search time 94.2821 Seconds

Perfect score: 794

Sequence: VLSSEGNEVOLVHWAKEYAD.....ALELFRKDIAAKYKELGYQG 153

1563.930 Million cell updates/sec

Title: US-09-455-978B-76

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 6019581 seqs, 976053577 residues

Total number of hits satisfying chosen parameters: 6019581

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%
Maximum Match 10%

Listing first 45 summaries

Database : Pending Patents-AA Main:*

1: /cgn2_6/picodata/2/paa/US06_COMB.pep:*

2: /cgn2_6/picodata/2/paa/US07_COMB.pep:*

3: /cgn2_6/picodata/2/paa/US08_COMB.pep:*

4: /cgn2_6/picodata/2/paa/US09_COMB.pep:*

5: /cgn2_6/picodata/2/paa/US081_COMB.pep:*

6: /cgn2_6/picodata/2/paa/US082_COMB.pep:*

7: /cgn2_6/picodata/2/paa/US083_COMB.pep:*

8: /cgn2_6/picodata/2/paa/US084_COMB.pep:*

9: /cgn2_6/picodata/2/paa/US085_COMB.pep:*

10: /cgn2_6/picodata/2/paa/US086_COMB.pep:*

11: /cgn2_6/picodata/2/paa/US087_COMB.pep:*

12: /cgn2_6/picodata/2/paa/US088_COMB.pep:*

13: /cgn2_6/picodata/2/paa/US089_COMB.pep:*

14: /cgn2_6/picodata/2/paa/US090_COMB.pep:*

15: /cgn2_6/picodata/2/paa/US091_COMB.pep:*

16: /cgn2_6/picodata/2/paa/US092_COMB.pep:*

17: /cgn2_6/picodata/2/paa/US093_COMB.pep:*

18: /cgn2_6/picodata/2/paa/US094_COMB.pep:*

19: /cgn2_6/picodata/2/paa/US095_COMB.pep:*

20: /cgn2_6/picodata/2/paa/US096_COMB.pep:*

21: /cgn2_6/picodata/2/paa/US097_COMB.pep:*

22: /cgn2_6/picodata/2/paa/US098_COMB.pep:*

23: /cgn2_6/picodata/2/paa/US099A_COMB.pep:*

24: /cgn2_6/picodata/2/paa/US099B_COMB.pep:*

25: /cgn2_6/picodata/2/paa/US100_COMB.pep:*

26: /cgn2_6/picodata/2/paa/US101_COMB.pep:*

27: /cgn2_6/picodata/2/paa/US102_COMB.pep:*

28: /cgn2_6/picodata/2/paa/US103_COMB.pep:*

29: /cgn2_6/picodata/2/paa/US104_COMB.pep:*

30: /cgn2_6/picodata/2/paa/US106_COMB.pep:*

31: /cgn2_6/picodata/2/paa/US117_COMB.pep:*

32: /cgn2_6/picodata/2/paa/US60_COMB.pep:*

33: /cgn2_6/picodata/2/paa/US60_COMB.pep:*

1 794 100.0 153 18 US-09-455-978B-76

2 794 100.0 153 21 US-09-791-537-51499

3 794 100.0 154 22 US-09-791-537-41444

4 789 99.4 153 22 US-09-791-537-23678

5 789 99.4 154 22 US-09-791-537-1857

6 788 99.2 154 22 US-09-791-537-25659

7 788 99.2 154 22 US-09-791-537-20088

8 786 99.0 154 22 US-09-791-537-23776

9 786 99.0 154 22 US-09-791-537-23776

10 786 99.0 154 22 US-09-791-537-104872

11 785 98.9 154 22 US-09-791-537-22335

12 785 98.9 154 22 US-09-791-537-23775

13 785 98.9 154 22 US-09-791-537-102151

14 785 98.9 154 22 US-09-791-537-102870

15 784 98.7 153 22 US-09-791-537-51498

16 784 98.7 154 22 US-09-791-537-18555

17 784 98.7 154 22 US-09-791-537-12322

18 784 98.7 154 22 US-09-791-537-22555

19 784 98.7 154 22 US-09-791-537-12870

20 783 98.6 151 22 US-09-791-537-71491

21 783 98.6 154 22 US-09-791-537-20794

22 783 98.6 154 22 US-09-791-537-3391

23 783 98.6 154 22 US-09-791-537-71339

24 783 98.6 154 22 US-09-791-537-102153

25 783 98.6 154 22 US-09-791-537-102916

26 783 98.6 154 22 US-09-791-537-151603

27 782 98.5 154 22 US-09-791-537-20811

28 782 98.5 154 22 US-09-791-537-7835

29 782 98.5 154 22 US-09-791-537-12896

30 781 98.4 154 22 US-09-791-537-1847

31 781 98.4 154 22 US-09-791-537-2671

32 781 98.4 154 22 US-09-791-537-12150

33 780 98.2 154 22 US-09-791-537-67801

34 780 98.2 154 22 US-09-791-537-12163

35 779 98.1 154 22 US-09-791-537-2158

36 779 98.1 154 22 US-09-791-537-24675

37 779 98.1 154 22 US-09-791-537-1099

38 778 98.0 151 22 US-09-791-537-18163

39 778 98.0 154 22 US-09-791-537-2159

40 778 98.0 154 22 US-09-791-537-24676

41 778 98.0 154 22 US-09-791-537-12159

42 775 97.6 154 22 US-09-791-537-20089

43 773 97.4 153 22 US-09-791-537-9391

44 772 97.2 154 22 US-09-791-537-2073

45 772 97.2 154 22 US-09-791-537-77660

ALIGNMENTS

RESULT 1 US-09-455-978B-76

; Application US/09455978B

; GENERAL INFORMATION:

; APPLICANT: Alam, Maqsudul

; ATTORNEY: Larsen, Randy

; TITLE OF INVENTION: HEMATE-HS AND HEMAT-BS AND THEIR USE IN MEDICINE AND MICROSENSORS

; FILE REFERENCE: 201040/1020

; CURRENT APPLICATION NUMBER: US/09/455,978B

; CURRENT FILING DATE: 1999-12-06

; NUMBER OF SEQ ID NOS: 86

; SOFTWARE: Patent ver. 2.1

; SEQ ID NO: 76

; LENGTH: 153

; TYPE: PRT

; ORGANISM: Sperm-whale myoglobin

US-09-455-978B-76

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	100.0%	100.0%	100.0%	DB 18;	Length 153;
2	100.0%	100.0%	100.0%	Pred. No. 2.4e-77;	
3	100.0%	100.0%	100.0%	Mismatches 0;	Indels 0; Gaps 0;

RESULT 2
US-09-700-708-4

; Sequence 4, Application US/09700708
; GENERAL INFORMATION:
; APPLICANT: ITAI, Aiko
; APPLICANT: ITAI, Reiko
; TITLE OF INVENTION: Method For Predicting Functions of Protein
; FILE REFERENCE: P20.94
; CURRENT APPLICATION NUMBER: US/09/700,708
; CURRENT FILING DATE: 2000-11-24
; PRIOR APPLICATION NUMBER: PCT/JP98/02302
; PRIOR FILING DATE: 1998-05-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Whale
; US-09-700-708-4

Query Match Score 794; DB 21; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.4e-77;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSEGENOLVLYWAKVEADVAHGQDILIRLPKSHPPTLEKEDRFKLKTEAEMKASED 60
Db 1 VLSEGENOLVLYWAKVEADVAHGQDILIRLPKSHPPTLEKEDRFKLKTEAEMKASED 60

Qy 61 LKGHGTVLTALGAILKKGGHRAELPKLAQSHATKHKIPIKYLEFISEAITHVLHSRHP 120
Db 61 LKGHGTVLTALGAILKKGGHRAELPKLAQSHATKHKIPIKYLEFISEAITHVLHSRHP 120

Qy 121 GDFGADAQAMNKALELFRKDIAKYKELGYQG 153
Db 121 GDFGADAQAMNKALELFRKDIAKYKELGYQG 153

Qy 121 GDFGADAQAMNKALELFRKDIAKYKELGYQG 153
Db 121 GDFGADAQAMNKALELFRKDIAKYKELGYQG 153

RESULT 4
US-09-791-537-41444

; Sequence 41444, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41444
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Phyleteter catodon
; US-09-791-537-41444

Query Match Score 789; DB 22; Length 153;
Best Local Similarity 99.4%; Pred. No. 8.5e-77;
Matches 152; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSEGENOLVLYWAKVEADVAHGQDILIRLPKSHPPTLEKEDRFKLKTEAEMKASED 60
Db 1 VLSEGENOLVLYWAKVEADVAHGQDILIRLPKSHPPTLEKEDRFKLKTEAEMKASED 60

Qy 61 LKGHGTVLTALGAILKKGGHRAELPKLAQSHATKHKIPIKYLEFISEAITHVLHSRHP 120
Db 61 LKGHGTVLTALGAILKKGGHRAELPKLAQSHATKHKIPIKYLEFISEAITHVLHSRHP 120

Qy 121 GDFGADAQAMNKALELFRKDIAKYKELGYQG 153
Db 121 GDFGADAQAMNKALELFRKDIAKYKELGYQG 153

RESULT 5
US-09-791-537-23678

; Sequence 23678, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23678
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Pdb 2MGK
; US-09-791-537-23678

Query Match Score 789; DB 22; Length 154;

Query Match Score 789; DB 22; Length 154;

Query Match Score 789; DB 22; Length 154;

Best Local Similarity 99.3%; Pred. No. 8.5e-77;
 Matches 152; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 99.2%; Score 788; DB 22; Length 154;
 Best Local Similarity 98.7%; Pred. No. 1.1e-76;
 Matches 151; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSSEGEGVQLVHWAKEYADVGAGHGDILIRLFKSHPETLEKFDRFKLKTEAEMKASED 60
 Db 2 VLSSEGEGVQLVHWAKEYADVGAGHGDILIRLFKSHPETLEKFDRFKLKTEAEMKASED 61

Qy 61 LKKHGTVTTLGAILKKKGHEAEELKPLAQSHATKHKIPKYLEFISEAIIHVLSRHP 120
 Db 62 LKKHGTVTTLGAILKKKGHEAEELKPLAQSHATKHKIPKYLEFISEAIIHVLSRHP 121

Qy 121 GDFGADAGAMMKALELFRKDIAAKYKELGYQG 153
 Db 122 GNFGADAGAMMKALELFRKDIAAKYKELGYQG 154

RESULT 6
 US-09-791-537-13857
 ; Sequence 13857, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 13857
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: pdb 1CIOA
 ; US-09-791-537-13857

Query Match 99.2%; Score 788; DB 22; Length 154;
 Best Local Similarity 98.7%; Pred. No. 1.1e-76;
 Matches 151; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSSEGEGVQLVHWAKEYADVGAGHGDILIRLFKSHPETLEKFDRFKLKTEAEMKASED 60
 Db 2 VLSSEGEGVQLVHWAKEYADVGAGHGDILIRLFKSHPETLEKFDRFKLKTEAEMKASED 61

Qy 61 LKKHGTVTTLGAILKKKGHEAEELKPLAQSHATKHKIPKYLEFISEAIIHVLSRHP 120
 Db 62 LKKHGTVTTLGAILKKKGHEAEELKPLAQSHATKHKIPKYLEFISEAIIHVLSRHP 121

Qy 121 GDFGADAGAMMKALELFRKDIAAKYKELGYQG 153
 Db 122 GNFGADAGAMMKALELFRKDIAAKYKELGYQG 154

RESULT 7
 US-09-791-537-22569
 ; Sequence 22569, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 22569
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: pdb 1MLM
 ; US-09-791-537-22569

RESULT 8
 US-09-791-537-23088
 ; Sequence 23088, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 23088
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: pdb 1MLR
 ; US-09-791-537-23088

Query Match 99.0%; Score 786; DB 22; Length 154;
 Best Local Similarity 98.7%; Pred. No. 1.8e-76;
 Matches 151; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSSEGEGVQLVHWAKEYADVGAGHGDILIRLFKSHPETLEKFDRFKLKTEAEMKASED 60
 Db 2 VLSSEGEGVQLVHWAKEYADVGAGHGDILIRLFKSHPETLEKFDRFKLKTEAEMKASED 61

Qy 61 LKKHGTVTTLGAILKKKGHEAEELKPLAQSHATKHKIPKYLEFISEAIIHVLSRHP 120
 Db 62 LKKHGTVTTLGAILKKKGHEAEELKPLAQSHATKHKIPKYLEFISEAIIHVLSRHP 121

Qy 121 GDFGADAGAMMKALELFRKDIAAKYKELGYQG 153
 Db 122 GNFGADAGAMMKALELFRKDIAAKYKELGYQG 154

RESULT 9
 US-09-791-537-23776
 ; Sequence 23776, Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 23776
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: pdb 2SP0
 ; US-09-791-537-23776

Query Match 99.0%; Score 786; DB 22; Length 154;
 Best Local Similarity 98.7%; Pred. No. 1.8e-76; Matches 151; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSGEWQLVHYAKVEADVAHGQDILIRLFKSHPTLEKDRFKLKTEAMKASED 60
 Db 2 VLSGEWQLVHYAKVEADVAHGQDILIRLFKSHPTLEKDRFKLKTEAMKASED 61

Qy 61 LKRGTVTITALGAILKKKGHHEAELPKLAQSHATKHPIKYLFISEAIIHVLSRHP 120
 Db 62 LKRGTVTITALGAILKKKGHHEAELPKLAQSHATKHPIKYLFISEAIIHVLSRHP 121

Qy 121 GDFGADAGMANKALELFRKDIAAKYKELGYQG 153
 Db 122 GNFGADAGMANKALELFRKDIAAKYKELGYQG 154

RESULT 10
 US-09-791-537-102872
 ; Sequence 102872 Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 102872
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: pdb 1CO9A
 US-09-791-537-102872

Query Match 99.0%; Score 786; DB 22; Length 154;
 Best Local Similarity 98.7%; Pred. No. 1.8e-76; Matches 151; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLSGEWQLVHYAKVEADVAHGQDILIRLFKSHPTLEKDRFKLKTEAMKASED 60
 Db 2 VLSGEWQLVHYAKVEADVAHGQDILIRLFKSHPTLEKDRFKLKTEAMKASED 61

Qy 61 LKRGTVTITALGAILKKKGHHEAELPKLAQSHATKHPIKYLFISEAIIHVLSRHP 120
 Db 62 LKRGTVTITALGAILKKKGHHEAELPKLAQSHATKHPIKYLFISEAIIHVLSRHP 121

Qy 121 GDFGADAGMANKALELFRKDIAAKYKELGYQG 153
 Db 122 GNFGADAGMANKALELFRKDIAAKYKELGYQG 154

RESULT 11
 US-09-791-537-22535
 ; Sequence 22535 Application US/09791537
 ; GENERAL INFORMATION:
 ; APPLICANT: Bionomix, Inc.
 ; APPLICANT: Debe, Derek
 ; APPLICANT: Danzer, Joseph
 ; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBER
 ; FILE REFERENCE: 261/210
 ; CURRENT APPLICATION NUMBER: US/09/791,537
 ; CURRENT FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 153055
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 22535
 ; LENGTH: 154
 ; TYPE: PRT
 ; ORGANISM: pdb 1MLF

; ORGANISM: pdb 1CH2A
US-09-791-537-102151

Query Match 98.9%; Score 785; DB 22; Length 154;
Best Local Similarity 98.7%; Pred. No. 2.3e-76;
Matches 151; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLSEGEWVLHWWAKEADYAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAEMKASED 60
Db 2 VLSEGEWVLHWWAKEADYAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAEMKASED 61

Qy 61 LKKHGTVTTLGAILKKGGHHEAELKPLAQSHATKHKIPKYLEFISEATHVLSRHP 120
Db 62 LKKHGTVTTLGAILKKGGHHEAELKPLAQSHATKHKIPKYLEFISEATHVLSRHP 121

Qy 121 GDFGADAGQAMNKALELFRKDIAAKYKELGYQG 153
Db 122 GNFGADAGQAMNKALELFRKDIAAKYKELGYQG 154

Qy 121 GDFGADAGQAMNKALELFRKDIAAKYKELGYQG 153
Db 121 GDFGADAGQAMNKALELFRKDIAAKYKELGYQG 153

RESULT 14
US-09-791-537-102898
; Sequence 102898, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE

; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153,055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 102898
; LENGTH: 154
; TYPE: PRT
; ORGANISM: pdb 1CP5A
US-09-791-537-102898

Query Match 98.9%; Score 785; DB 22; Length 154;
Best Local Similarity 98.7%; Pred. No. 2.3e-76;
Matches 151; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VLSEGEWVLHWWAKEADYAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAEMKASED 60
Db 2 VLSEGEWVLHWWAKEADYAGHGQDILIRLFKSHPETLEKFDRFKHLKTEAEMKASED 61

Qy 61 LKKHGTVTTLGAILKKGGHHEAELKPLAQSHATKHKIPKYLEFISEATHVLSRHP 120
Db 62 LKKHGTVTTLGAILKKGGHHEAELKPLAQSHATKHKIPKYLEFISEATHVLSRHP 121

Qy 121 GDFGADAGQAMNKALELFRKDIAAKYKELGYQG 153
Db 122 GNFGADAGQAMNKALELFRKDIAAKYKELGYQG 154

RESULT 15
US-09-791-537-51498
; Sequence 51498, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; NUMBER OF SEQ ID NOS: 153,055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51498
; LENGTH: 153

Search completed: August 10, 2004, 15:39:00
Job time : 95.2821 secs

B
l
a
n
k

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Inc.

OM protein - protein search, using sw model

Run on: August 10, 2004, 15:26:56 ; Search time 12.04 Seconds

Title: US-09-455-978B-76

Perfect score: 794

Sequence: 1 VLSSEGQWQLVHWAKEBAD.....ALLEFRKDIAAKYKELGYQG 153

Scoring table: BLOSUM62

Gapext 0.5

Searched: 618821 seqs, 107722064 residues

Total number of hits satisfying chosen parameters: 618821

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA New: *

1: /cgn2_6/picodata/2/paa/PCT_NEW_COMB_pep:*

2: /cgn2_6/picodata/2/paa/US05_NEW_COMB_pep:*

3: /cgn2_6/picodata/2/paa/US07_NEW_COMB_pep:*

4: /cgn2_6/picodata/2/paa/US08_NEW_COMB_pep:*

5: /cgn2_6/picodata/2/paa/US10_NEW_COMB_pep:*

6: /cgn2_6/picodata/2/paa/US10_NEW_COMB_pep:*

7: /cgn2_6/picodata/2/paa/US60_NEW_COMB_pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match %	Length	DB ID	Description
1	69.9	88.0	153	6	US-10-784-004-632	Sequence 632, APP
2	69.9	88.0	153	6	US-10-784-004-1039	Sequence 1039, APP
3	69.9	88.0	154	6	US-10-170-205E-32946	Sequence 33330, A
4	69.9	88.0	154	6	US-10-170-205E-33330	Sequence 33331, A
5	69.9	88.0	154	6	US-10-784-004-310	Sequence 310, APP
6	64.7	81.5	154	6	US-10-784-004-898	Sequence 898, APP
7	47.2	59.4	113	6	US-10-793-47-4103	Sequence 4103, APP
8	32.7	41.2	79	6	US-10-425-115-279565	Sequence 279565, Sequence 4102, APP
9	31.4	39.5	71	6	US-10-793-47-4102	Sequence 750, APP
10	21.2	26.8	190	1	PCT-US04-07412-750	Sequence 21, APP
11	21.2	26.8	190	1	PCT-US02-17443-214	Sequence 216, APP
12	21.2	26.8	190	1	PCT-US02-17443-216	Sequence 750, APP
13	21.2	26.8	190	6	US-10-389-553-750	Sequence 19233, A
14	21.2	26.8	190	6	US-10-170-205E-19233	Sequence 7507, APP
15	21.2	26.8	190	6	US-10-170-205E-19233	Sequence 21, APP
16	15.9	20.0	142	6	US-10-776-172-21	Sequence 214, APP
17	14.0	17.6	141	6	US-10-776-172-21	Sequence 19, APP
18	14.0	17.6	141	6	US-10-897-005-28	Sequence 26, APP
19	12.8	16.1	141	6	US-10-776-172-19	Sequence 29, APP
20	12.8	16.1	141	6	US-10-897-005-26	Sequence 29, APP
21	12.7	16.0	146	6	US-10-776-172-22	Sequence 23524, A
22	12.7	16.0	146	6	US-10-897-005-29	Sequence 903, APP
23	12.5	15.7	147	6	US-10-170-205E-23524	Sequence 147, APP
24	12.5	15.7	147	7	US-60-563-440-970	Sequence 20561, APP
25	12.4	15.6	120	6	US-10-125-115-20641	Sequence 201, APP
26	15.6	147	1	PCT-US03-36002-201	Sequence 1593, APP	

ALIGNMENTS

RESULT 1
US-10-784-004-632
; Sequence 632, Application US/10784004
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201_6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 632
; LENGTH: 153
; TYPE: PRT
; ORGANISM: human
US-10-784-004-632

Query Match 88.0%; Score 699; DB 6; Length 153;
Best Local Similarity 84.9%; Pred. No. 4.7e-4;
Matches 129; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

RESULT 2
US-10-784-004-1039
; Sequence 1039, Application US/10784004
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201_6029-00000
; CURRENT APPLICATION NUMBER: US/10/784,004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 1039
; LENGTH: 153
; TYPE: PRT
; ORGANISM: human

US-10-784-004-1039

Query Match 88.0% ; Score 699; DB 6; Length 153;
Best Local Similarity 84.9%; Pred. No. 4.7e-64;
Matches 129; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

Qy 2 LSEGEWOLVLYWAKVEADVAGHQDILIRLFKSHPTELEKDFRKHLKTEAEMKASEDL 61
Db 1 LSDEGWOLVLYWAKVEADVAGHQDILIRLFKSHPTELEKDFRKHLKTEAEMKASEDL 62

Qy 62 KKHGTVTITALGAILKKKGHEAELKPAQSHATKHKPIKYLEFISBAITHVLHSRHPG 121
Db 63 KKHGATVLTALGG3LKKKGHEAELKPAQSHATKHKPIKYLEFISCCIIQVLQSXHPG 122

Qy 122 DFGADAAQAMNKALELFRKDIAAKYKELGYQG 153
Db 123 DFGADAAQAMNKALELFRKDMSNYKELGFQG 154

RESULT 5
US-10-170-205E-33331
; Sequence 33331, Application US/10170205E

; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170, 205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SEQ ID NO: 33331
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-33331

Query Match 88.0% ; Score 699; DB 6; Length 154;
Best Local Similarity 84.9%; Pred. No. 4.7e-64;
Matches 129; Conservative 13; Mismatches 10; Indels 0; Gaps 0;

Qy 2 LSEGEWOLVLYWAKVEADVAGHQDILIRLFKSHPTELEKDFRKHLKTEAEMKASEDL 61
Db 3 LSDEGWOLVLYWAKVEADVAGHQDILIRLFKSHPTELEKDFRKHLKTEAEMKASEDL 62

Qy 62 KKHGTVTITALGAILKKKGHEAELKPAQSHATKHKPIKYLEFISBAITHVLHSRHPG 121
Db 63 KKHGATVLTALGG3LKKKGHEAELKPAQSHATKHKPIKYLEFISCCIIQVLQSXHPG 122

Qy 122 DFGADAAQAMNKALELFRKDIAAKYKELGYQG 153
Db 123 DFGADAAQAMNKALELFRKDMSNYKELGFQG 154

RESULT 6
US-10-784-004-310
; Sequence 310, Application US/10784004

; GENERAL INFORMATION:
; APPLICANT: Biogen Idec
; TITLE OF INVENTION: Surrogate Markers of Pain
; FILE REFERENCE: 08201.6029-00000
; CURRENT APPLICATION NUMBER: US/10/784, 004
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 1251
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 310
; LENGTH: 154
; TYPE: PRT
; ORGANISM: rat
US-10-784-004-310

Query Match 81.5% ; Score 647; DB 6; Length 154;
Best Local Similarity 78.9%; Pred. No. 1e-58;
Matches 120; Conservative 18; Mismatches 14; Indels 0; Gaps 0;

Qy 2 LSEGEWOLVLYWAKVEADVAGHQDILIRLFKSHPTELEKDFRKHLKTEAEMKASEDL 61
; Sequence Match

NAME/KEY: UNSURE
 LOCATION: 84
 OTHER INFORMATION: Xaa=Phe or Ile or Leu or Met or Val

NAME/KEY: UNSURE
 LOCATION: 95
 OTHER INFORMATION: Xaa=Pro or Thr
 US-10-793-479-4103

Query Match 59.4%; Score 472; DB 6; Length 113;
 Best Local Similarity 85.1%; Pred. No. 6.3e-41;
 Matches 86; Indels 0; Gaps 0;

Qy 2 LSEGENOLVLIHWAKVEADVAGHGDIDLIRLFKSHPETLEKFDKFHLKTEAEMKASEDL 61
 Db 3 LSDGENQMVNLWKGKVBDLGHGQEVLISLKAHETLEKFDKFHLKSEDEMKAESDL 62

Query Match 59.4%; Score 472; DB 6; Length 113;
 Best Local Similarity 85.1%; Pred. No. 6.3e-41;
 Matches 86; Indels 0; Gaps 0;

Qy 2 LSEGENOLVLIHWAKVEADVAGHGDIDLIRLFKSHPETLEKFDKFHLKTEAEMKASEDL 61
 Db 3 LSDGENQMVNLWKGKVBDLGHGQEVLISLKAHETLEKFDKFHLKSEDEMKAESDL 62

RESULT 9
 US 10-425-115-279565
 Sequence 279565, Application US/10425115
 GENERAL INFORMATION:
 APPLICANT: Biogen Idec
 TITLE OF INVENTION: Surrogate Markers of Pain
 FILE REFERENCE: 06201.6029-00000
 CURRENT APPLICATION NUMBER: US/10/784,004
 CURRENT FILING DATE: 2004-02-20
 NUMBER OF SEQ ID NOS: 1251
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 893
 LENGTH: 154
 TYPE: PRT
 ORGANISM: rat
 US-10-784-004-8948

Query Match 81.5%; Score 647; DB 6; Length 154;
 Best Local Similarity 78.9%; Pred. No. 1e-58;
 Matches 18; Mismatches 14; Indels 0; Gaps 0;

Qy 2 LSEGENOLVLIHWAKVEADVAGHGDIDLIRLFKSHPETLEKFDKFHLKTEAEMKASEDL 61
 Db 3 LSDGENQMVNLWKGKVBDLGHGQEVLISLKAHETLEKFDKFHLKSEDEMKAESDL 62

Query Match 81.5%; Score 647; DB 6; Length 154;
 Best Local Similarity 78.9%; Pred. No. 1e-58;
 Matches 18; Mismatches 14; Indels 0; Gaps 0;

Qy 2 LSEGENOLVLIHWAKVEADVAGHGDIDLIRLFKSHPETLEKFDKFHLKTEAEMKASEDL 61
 Db 3 LSDGENQMVNLWKGKVBDLGHGQEVLISLKAHETLEKFDKFHLKSEDEMKAESDL 62

Query Match 81.5%; Score 647; DB 6; Length 154;
 Best Local Similarity 78.9%; Pred. No. 1e-58;
 Matches 18; Mismatches 14; Indels 0; Gaps 0;

Qy 2 LSEGENOLVLIHWAKVEADVAGHGDIDLIRLFKSHPETLEKFDKFHLKTEAEMKASEDL 61
 Db 3 LSDGENQMVNLWKGKVBDLGHGQEVLISLKAHETLEKFDKFHLKSEDEMKAESDL 62

RESULT 10
 US-10-793-479-4102
 Sequence 4102, Application US/10793479
 GENERAL INFORMATION:
 APPLICANT: Dumas Milne Edwards, J.B.
 APPLICANT: Ducleir, A.
 APPLICANT: Giordano, J.Y.
 TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
 FILE REFERENCE: 59.US2.REG
 CURRENT APPLICATION NUMBER: US/10/793,479
 PRIOR APPLICATION NUMBER: US/09/513,999
 PRIOR FILING DATE: 2004-03-03
 PRIOR FILING DATE: 2000-02-24
 NUMBER OF SEQ ID NOS: 36681
 SOFTWARE: Patent-pm
 SEQ ID NO 4103
 LENGTH: 113
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE: SIGNAL
 LOCATION: -20..-1
 OTHER INFORMATION: score 3.5
 OTHER INFORMATION: seq WQLVNWKGVEA/DI
 FEATURE:

SOFTWARE: Patent.pml
SEQ ID NO: 4102
LENGTH: 71
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE: SIGNAL
NAME/KEY: SIGNAL
LOCATION: -20 .. -1
OTHER INFORMATION: score 3.5
OTHER INFORMATION: seq WQLVILNVWGKVEA/DI
US-10-793-479-4102

Query Match 39.5%; Score 314; DB 6; Length 71;
Best Local Similarity 82.6%; Pred. No. 5.7e-5; Indels 0; Gaps 0;
Matches 57; Conservative 7; Mismatches 5; Gaps 0;

Qy 2 LSEGEGWOLVLYWAKVEADVAGHGGDQDLIRLFKSHPTELEKFDRFKHLKTEAEMKASEDL 61
Db 3 LSDGEWQVLNVWGKVEADPGHQEVILRLFKGHPPTELEKFDFKHLKSDDEMKAESDL 62

Qy 62 KKHGTVLTT 70
Db 63 KKHGATVLT 71

RESULT 11
PCT-US04-07412-750
Sequence 750, Application PC/TUS0407412
GENERAL INFORMATION
APPLICANT: Tang, Y. Tom
APPLICANT: Wang, Jian-rui
APPLICANT: Ben, Feiyan
APPLICANT: Zhou, Fing
APPLICANT: Ma, Yurqing
APPLICANT: Ghosh, Malabika
APPLICANT: Xue, Aidong J.
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Dunrui
APPLICANT: Goodrich, Ryle W.
APPLICANT: Chen, Rui-hong
APPLICANT: Wehrman, Tom
APPLICANT: Weng, Gezhi
APPLICANT: Wang, Zhiwei
APPLICANT: Boyle, Bryan J.
APPLICANT: Drmana, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 822CIP/BCT

CURRENT FILING DATE: 2004-03-19
PRIORITY APPLICATION NUMBER: US 10/389,559
PRIOR FILING DATE: 2003-03-14
PRIOR APPLICATION NUMBER: US 60/365,264
PRIOR FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: US 60/340,187
PRIOR FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 10/296,115
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: PCT/US00/35017
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 10/275,027
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1920

SOFTWARE: pt_FL_genes Version 6.0
SEQ ID NO 750
LENGTH: 190
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US04-07412-750

Query Match 26.8%; Score 212.5; DB 1; Length 190;
Best Local Similarity 31.6%; Pred. No. 5.4e-14;
Matches 49; Conservative 31; Mismatches 68; Indels 7; Gaps 2;

Qy 2 LSEGEWQVLVLYWAKVEADVAGHGGDQDLIRLFKSHPTELEKFDRFKHLKTEAEMKASEDL 61
Db 19 LSEARECAVQANWARYANCEDGVAVLVRFFNPFSAKOFSPQKFMEDEPLMERSPQL 78

Qy 62 KKHGTVLTTALGAILKKKGHHAE----LKPLAQSHATKHKPIKYLEFISBAAIHYVLH 116
Db 79 RKHACRUMGALNTVNL--HPDKVSVLALVGKAHALKRVPEVYFKILSGVILEVVA 136

Qy 117 SRHPGDFGADAOGAMMKALEFRKDIAAKYKELGY 151
Db 137 ERFASDFPPETORAWAKLRGLIYSHVTAAKVEVGW 171

RESULT 12
PCT-US02-17443-214
Sequence 214, Application PC/TUS0217443

GENERAL INFORMATION:
APPLICANT: Curagen Corporation
APPLICANT: Agee, Michele
APPLICANT: Alsobrook, John
APPLICANT: Anderson, David
APPLICANT: Bergs, Constance
APPLICANT: Burgess, Catherine
APPLICANT: Casman, Stacie
APPLICANT: Catterton, Elina
APPLICANT: DiPippo, Vincent
APPLICANT: Edinger, Shlomit
APPLICANT: Eisen, Andrew
APPLICANT: Ellerman, Karen
APPLICANT: Gangoli, Esha
APPLICANT: Gerlach, Valerie
APPLICANT: Goldman, Linda
APPLICANT: Guo, Xiaoja
APPLICANT: Herrmann, John
APPLICANT: Hjalt, Tord
APPLICANT: Ji, Weizhen
APPLICANT: Kekuda, Ramesh
APPLICANT: Khrantsov, Nikolai
APPLICANT: Lepley, Denise
APPLICANT: Li, Li
APPLICANT: Liu, Xiaohong
APPLICANT: Malivankar, Uriel
APPLICANT: Miller, Charles
APPLICANT: Milliet, Isabelle
APPLICANT: Ort, Tatiana
APPLICANT: Patilguru, Muraldhara
APPLICANT: Patterson, Meera
APPLICANT: Pena, Carol
APPLICANT: Rastelli, Luca
APPLICANT: Rieger, Daniel
APPLICANT: Rothemberg, Mark
APPLICANT: Shenoy, Suresh
APPLICANT: Shimkets, Richard
APPLICANT: Smithson, Glenna
APPLICANT: Spaderna, Steven
APPLICANT: Spytek, Kimberly
APPLICANT: Stone, David
APPLICANT: Vernet, Corinne
APPLICANT: Zhong, Haibong
APPLICANT: Zhong, Mei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-3777C-061

CURRENT APPLICATION NUMBER: PCT/US02/17443
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 10/160, 619
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295, 661
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/295, 607
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/296, 404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296, 418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296, 575
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/297, 414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297, 573
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/297, 567
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/298, 285
; PRIOR FILING DATE: 2001-06-14
; REMAINING Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 539
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO: 214
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT/US02/17443-214

Query Match 26.8%; Score 212.5; DB 1; Length 190;
Best Local Similarity 31.4%; Pred. No. 5.4e-14;
Matches 49; Conservative 31; Mismatches 68; Indels 7; Gaps 2;

Qy 2 LSEGEQVLVHWAKEYADWAGHQDQLLIRLFLFKSHPETLEKFDRFKLKTCADMKAESDL 61
Db 19 LSEAERKAVQAMWARYANCEDVGFNFSQAKQYFSQFRKHMDFPLEMERSPQL 78

Qy 62 KKHGIVTVLTALGAILEKKKGHEAE---LPLAQSHATHKHKIPRKYLFSEATHVHLA 116
Db 79 RKHACRVMGALNTVVNL-HDPDKVSSVALVGKAHALHKVEPVYFKLILGVILLEVVA 136

Qy 117 SRHPGDFGADQGAMMKALLEPLRDKIAKYKELGY 151
Db 137 EEFASDFPPETQRAWAKLRLGLIYSHVTAAYKEVGW 171

RESULT 13
PCT-US02-17443-216
; Sequence 216, Application PC/TUS02/17443
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation
; APPLICANT: Agee, Michele
; APPLICANT: Alsobrook, John
; APPLICANT: Anderson, David
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Catterton, Elina
; APPLICANT: DiPippo, Vincent
; APPLICANT: Edinger, Shlomit
; APPLICANT: Eisen, Andrew
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha
; APPLICANT: Gerlach, Valerie
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Herrmann, John
; APPLICANT: Hjalt, Tord
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Khrantsov, Nikolai

Query Match 26.8%; Score 212.5; DB 1; Length 190;
Best Local Similarity 31.6%; Pred. No. 5.4e-14;
Matches 49; Conservative 31; Mismatches 68; Indels 7; Gaps 2;

Qy 2 LSEGEMQVLHVWAKYEADVAGHGQDILIRLFLKFSPTELKFDRPKHLKTEAEMKAESDL 61
Db 19 LSEAERKAVQAMWARYANCEDVGFVYLRFVNPFSAKOYFSQFRKHMDFPLEMERSPQL 78

Qy 62 KKHGIVTVLTALGAILEKKKGHEAE---LKPLAQSHATHKHKIPRKYLFSEATHVHLH 116
Db 79 RKHACRVMGALNTVVNL-HDPDKVSSVALVGKAHALHKVEPVYFKLILGVILLEVVA 136

Qy 117 SRHPGDFGADQGAMMKALLEPLRDKIAKYKELGY 151
Db 137 EEFASDFPPETQRAWAKLRLGLIYSHVTAAYKEVGW 171

SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO: 216
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-17443-216

Query Match 26.8%; Score 212.5; DB 1; Length 190;
Best Local Similarity 31.6%; Pred. No. 5.4e-14;
Matches 49; Conservative 31; Mismatches 68; Indels 7; Gaps 2;

Qy 2 LSEGEMQVLHVWAKYEADVAGHGQDILIRLFLKFSPTELKFDRPKHLKTEAEMKAESDL 61
Db 19 LSEAERKAVQAMWARYANCEDVGFVYLRFVNPFSAKOYFSQFRKHMDFPLEMERSPQL 78

Qy 62 KKHGIVTVLTALGAILEKKKGHEAE---LKPLAQSHATHKHKIPRKYLFSEATHVHLH 116
Db 79 RKHACRVMGALNTVVNL-HDPDKVSSVALVGKAHALHKVEPVYFKLILGVILLEVVA 136

Qy 117 SRHPGDFGADQGAMMKALLEPLRDKIAKYKELGY 151
Db 137 EEFASDFPPETQRAWAKLRLGLIYSHVTAAYKEVGW 171

US-10-389-559-750 ; Sequence 750, Application US/10389559
 GENERAL INFORMATION: Application US/10389559
 APPLICANT: Tang, Y. Tom
 APPLICANT: Wang, Jian-rui
 APPLICANT: Zhang, Jie
 APPLICANT: Ren, Feiyan
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yurong
 APPLICANT: Ghosh, Malabika
 APPLICANT: Xue, Aidong J.
 APPLICANT: Asundi, Vinod
 APPLICANT: Zhao, Qing A.
 APPLICANT: Wang, Dunrui
 APPLICANT: Goodrich, Rytle W.
 APPLICANT: Chen, Rui-hong
 APPLICANT: Wehrman, Tom
 APPLICANT: Weng, Gezhi
 APPLICANT: Wang, Zhiwei
 APPLICANT: Boyle, Bryan J.
 APPLICANT: Drmanac, Radmila T.
 TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
 FILE REFERENCE: 8222
 CURRENT APPLICATION NUMBER: US/10/389,559
 CURRENT FILING DATE: 2003-03-14
 PRIOR APPLICATION NUMBER: US 60/365,264
 PRIOR FILING DATE: 2002-03-14
 PRIOR APPLICATION NUMBER: US 60/340,187
 PRIOR FILING DATE: 2002-12-12
 PRIOR APPLICATION NUMBER: PCT/US00/35017
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: US 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: US 09/488,725
 PRIOR FILING DATE: 2000-01-21
 PRIOR APPLICATION NUMBER: PCT/US01/02643
 PRIOR FILING DATE: 2001-01-25
 PRIOR APPLICATION NUMBER: US 09/491,404
 PRIOR FILING DATE: 2000-01-25
 PRIOR APPLICATION NUMBER: PCT/US01/03880
 PRIOR FILING DATE: 2001-02-05
 PRIOR APPLICATION NUMBER: US 09/560,875
 PRIOR FILING DATE: 2000-04-27
 PRIOR APPLICATION NUMBER: US 09/496,914
 PRIOR FILING DATE: 2000-02-03
 Remaining Prior Application data removed - See File Wrapper or PALM.
 SOFTWARE: pt_FL SEQ ID NO: 1920
 SEQ ID NO: 750 LENGTH: 190
 TYPE: PT
 ORGANISM: Homo sapiens Version 6.0
 US-10-389-559-750

Query Match 26.8%; Score 212.5; DB 6; Length 190;
 Best Local Similarity 31.6%; Pred. No. 5.4e-14;
 Matches 49; Conservative 31; Mismatches 68; Indels 7; Gaps 2;

Qy 2 LSEGEWQLVLYWAKVEADVGHQDILRLFKSHPETLEKDRPKHLKTEAMKASEDL 61
 Db 19 LSEAERKAVQAMARLYANCEDGVAILVRFVNFPSSAKQYFSQFKHMDPELMERSPQL 78
 Qy 62 KKHGTVTTLGAILKKKGHHBAA---LKPLAQSHATKHKIPKYLEFISBAIHWLH 116
 Db 79 RKAACRVMGAAANTVNL-HDPDKVSSVLAIVGKAHLKERVEPYFKLISGVILLEVVA 136
 Qy 117 SRHPGDFGADAQGMNKAELFRKDIAAKYKELGY 151
 Db 137 EEFASDFPPETQRAWKLRLGIYSHVTAYKEVGW 171

Search completed: August 10, 2004, 15:40:16
 Job time : 13.04 secs

; Sequence 19233, Application US/10170205E
 ; GENERAL INFORMATION:
 ; APPLICANT: ADAMS, Mark
 ; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
 ; FILE REFERENCE: CJO01381
 ; CURRENT APPLICATION NUMBER: US/10/170,205E
 ; CURRENT FILING DATE: 2002-06-13
 ; NUMBER OF SEQ ID NOS: 49312
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO: 19233
 ; LENGTH: 190
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-170-205E-19233
 Query Match 26.8%; Score 212.5; DB 6; Length 190;
 Best Local Similarity 31.6%; Pred. No. 5.4e-14;
 Matches 49; Conservative 31; Mismatches 68; Indels 7; Gaps 2;
 Qy 2 LSEGEWQLVLYWAKVEADVGHQDILRLFKSHPETLEKDRPKHLKTEAMKASEDL 61
 Db 19 LSEAERKAVQAMARLYANCEDGVAILVRFVNFPSSAKQYFSQFKHMDPELMERSPQL 78
 Qy 62 KKHGTVTTLGAILKKKGHHBAA---LKPLAQSHATKHKIPKYLEFISBAIHWLH 116
 Db 79 RKAACRVMGAAANTVNL-HDPDKVSSVLAIVGKAHLKERVEPYFKLISGVILLEVVA 136
 Qy 117 SRHPGDFGADAQGMNKAELFRKDIAAKYKELGY 151
 Db 137 EEFASDFPPETQRAWKLRLGIYSHVTAYKEVGW 171

Best Local Similarity 100.0%; Pred. No. 8.4e-59;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 1 VLSSEGKQVLIHVWAKVEADYAGHGQDILIRLFLKSHPBTLEKFDRFKLKTEAEMKASED 60
 Db 1 VLSSEGKQVLIHVWAKVEADYAGHGQDILIRLFLKSHPBTLEKFDRFKLKTEAEMKASED 60

Query 61 LKKHGTIVLTALGAILLKKGHHAEELKPLAQOSHATKHPIKYLEFISEATHVLHSRHP 120
 Db 61 LKKHGTIVLTALGAILLKKGHHAEELKPLAQOSHATKHPIKYLEFISEATHVLHSRHP 120

Query 121 GDFGADAGAMNKALEFRKDIAAKYKELGYQ 153
 Db 121 GDFGADAGAMNKALEFRKDIAAKYKELGYQ 153

RESULT 2

MYWHW

myoglobin - dwarf sperm whale

C:Species: *Mogia simus* (dwarf sperm whale)

C:Accession: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 03-Mar-2000

C:Keywords: chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier

C:Superfamily: Globin homology

C:Domain: Globin homology

F:64/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:93/Binding site: heme iron (His) (proximal axial ligand) #status predicted

RESULT 4

MYWHH

myoglobin - humpback whale

C:Species: *Megaptera novaeangliae* (humpback whale)

C:Accession: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 03-Mar-2000

C:Keywords: chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier

C:Superfamily: Globin homology

C:Domain: Globin homology

F:64/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:93/Binding site: heme iron (His) (proximal axial ligand) #status predicted

RESULT 5

MYDDAR

myoglobin - Amazon dolphin

C:Species: *Inia geoffrensis* (Amazon dolphin)

C:Accession: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 03-Mar-2000

C:Keywords: chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier

C:Superfamily: Globin homology

C:Domain: Globin homology

F:64/Binding site: oxygen (His) (distal axial ligand) #status predicted

F:93/Binding site: heme iron (His) (proximal axial ligand) #status predicted

A;Residues: 1-153 <DWU>	
C;Comment: This myoglobin was isolated from skeletal muscle.	
C;SuperFamily: Globin; globin homology	
C;Keywords: chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier	
F;2-147/Domain: Globin homology <GLB>	
F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted	
F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted	
Query Match	92.9%; Score 738; DB 1; Length 153;
Best Local Similarity	91.4%; Pred. No. 3.6e-54;
Matches	139; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
Qy	2 LSEGEWQLVHLWWKVEADYAGHGDILRLFKSHPTELEKFDFKHLKTEAMKASEDL 61
Db	2 LSDGEWQLVNLNIGKVEADLGLAGQDVLIRLFKSHPTELEKFDFKHLKTEAMKASEDL 61
Qy	62 KKGGTYVTALGAILKKKGHEAEELKPLAQSHATRKHKPIKYLFISEAIIHVLSRHPG 121
Db	62 KKGGNTVTALGAILKKKGHEAEELKPLAQSHATRKHKPIKYLFISEAIIHVLSRHPG 121
Qy	122 DFGADAQGMNKKALELFRKDIAAKYKELGQY 153
Db	122 DFGADAQGMNKKALELFRKDIAAKYKELGFGH 153
RESULT 6	
MYWHIT	
myoglobin - pilot whale	
C;Species: Globicephala melasena (pilot whale)	
C;Database: 27-Nov-1985 #sequence revision 27-Nov-1985 #text_change 03-Mar-2000	
C;Accession: A02496	
R;Jones, B.N.; Duwlet, F.E.; Lehman, L.D.; Garner, M.H.; Bogardt Jr., R.A.; Garner, W.H.	
Biochemistry 17, 1971-1974, 1978	
F;2-147/Domain: Globin homology <GLB>	
F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted	
F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted	
Query Match	92.8%; Score 737; DB 1; Length 153;
Best Local Similarity	92.1%; Pred. No. 4.3e-54;
Matches	140; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
Qy	2 LSEGEWQLVHLWWKVEADYAGHGDILRLFKSHPTELEKFDFKHLKTEAMKASEDL 61
Db	2 LSDGEWQLVNLNIGKVEADLGLAGQDVLIRLFKSHPTELEKFDFKHLKTEAMKASEDL 61
Qy	62 KKGGTYVTALGAILKKKGHEAEELKPLAQSHATRKHKPIKYLFISEAIIHVLSRHPG 121
Db	62 KKGGNTVTALGAILKKKGHEAEELKPLAQSHATRKHKPIKYLFISEAIIHVLSRHPG 121
Qy	122 DFGADAQGMNKKALELFRKDIAAKYKELGQY 153
Db	122 EFGADAQGMNKKALELFRKDIAAKYKELGFGH 153
RESULT 7	
MYWHL	
myoglobin - killer whale	
C;Species: Orcinus orca (killer whale)	
C;Database: 27-Nov-1985 #sequence revision 27-Nov-1985 #text_change 03-Mar-2000	
C;Accession: A92956; A90614; A02495	
R;Neuth, J.L.; Jones, B.N.; Gurd, F.R.N.	
J. Mol. Evol. 17, 163-166, 1981	
F;2-147/Domain: Globin homology <GLB>	
F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted	
F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted	
Query Match	92.8%; Score 737; DB 1; Length 153;
Best Local Similarity	92.1%; Pred. No. 4.3e-54;
Matches	140; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
Qy	2 LSEGEWQLVHLWWKVEADYAGHGDILRLFKSHPTELEKFDFKHLKTEAMKASEDL 61
Db	2 LSDGEWQLVNLNIGKVEADLGLAGQDVLIRLFKSHPTELEKFDFKHLKTEAMKASEDL 61
Qy	62 KKGGTYVTALGAILKKKGHEAEELKPLAQSHATRKHKPIKYLFISEAIIHVLSRHPG 121
Db	62 KKGGNTVTALGAILKKKGHEAEELKPLAQSHATRKHKPIKYLFISEAIIHVLSRHPG 121
Qy	122 DFGADAQGMNKKALELFRKDIAAKYKELGQY 153
Db	122 EFGADAQGMNKKALELFRKDIAAKYKELGFGH 153

A;Note: the saddleback dolphin is also called Black Sea dolphin
 C;Date: 27-Nov-1985 #sequence_revision 31-Mar-1991 #text_change 03-Mar-2000
 C;Accession: A2230; A2509; A37510; A37509; A37510
 R;Wang, C.C.; Avilla, R.; Jones, B.N.; Gurd, F.R.N.
 Biochemistry 16, 4978-4981, 1977
 A;Title: Primary structure of the major component myoglobin of Pacific common dolphin
 A;Reference number: A26230; MUID: 911808
 A;Accession: A2230
 A;Molecule type: protein
 A;Residues: 1-153 <WAN>
 R;Klub, I.; Bakardjieva, A.
 PBS Lett. 17, 31-34, 1971
 A;Title: Primary structure of N-terminal part of molecule of dolphin myoglobin.
 A;Reference number: A02497
 A;Accession: A02497
 A;Molecule type: protein
 A;Residues: 1-20, 'V', 22-27, 'I', 29-31 <KLJ>
 R;Kazadzova, M.; Nedkov, P.; Bakardjieva, A.; Genov, N.
 Biochim. Biophys. Acta 221, 136-139, 1970
 A;Title: Difference in amino acid sequence between dolphin and sperm whale myoglobins.
 A;Reference number: A37509; MUID: 71014229; PMID: 5473803
 A;Accession: A37509
 A;Molecule type: protein
 A;Residues: 31-65, 'D', 67-121, 'Q', 123-153 <KAR>
 A;Note: the amidation states of residues 60, 85, and 132 were not determined and were assumed to be the same, December 1977
 A;Reference number: A37510
 A;Contents: revision to residue 26
 A;Accession: A37510
 A;Molecule type: protein
 A;Residues: 26 <KL2>
 A;Superfamily: globin homology
 C;Keywords: chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier
 F;2-147/Domain: globin homology <GLB>
 F;2-147/Binding site: oxygen (His) (distal axial ligand) #status predicted
 F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted
 Query Match 92.3%; Score 733; DB 1; Length 153;
 Best Local Similarity 90.8%; Pred. No. 9..3e-5%; Indels 0; Gaps 0;
 Matches 138; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 2 LSEGEWQLVLHWAKEADVAGHQDILIRLFKSHPETLEKDFRKHKTAEAKMASEDL 61
 Db 2 LSEGEWQLVLNNGKVEDLAGHQDILIRLFKSHPETLEKDFRKHKTAEAKMASEDL 61
 Qy 62 KKHGTVTTLAIGAILKKGHHEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPG 121
 Db 62 KKHGTVTTLAIGAILKKGHHEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPA 121
 Qy 122 DFGADAQAMMNKAELFRKDIAAKYKELGYQ 153
 Db 122 EFGADAQAMMNKAELFRKDIAAKYKELGFHG 153

RESULT 10
 MWPB
 myoglobin - harbor porpoise
 C;Species: Phocoena phocoena (harbor porpoise)
 C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 03-Mar-2000
 C;Accession: B92045; A90415; A02498
 R;bradshaw, R.A.; Gurd, F.R.N.
 J.Biol.Chem. 244, 2467-2481, 1969
 A;Title: Comparison of myoglobins from harbor seal, porpoise, and sperm whale.
 A;Reference number: A9045; MUID: 69177451; PMID: 5788005

A;Accession: B92045
 A;Molecule type: protein
 A;Residues: 1-82, EAN, 86-153 <BRA>
 A;Note: this sequence has been revised in reference A90415
 A;Note: this is the final paper in a series
 R;Meuth, J.L.; Jones, B.N.; Garner, W.H.; Gurd, F.R.N.
 Biochemistry 17, 3429-3431, 1978
 A;Title: Complete amino acid sequence of the myoglobin from the Dall porpoise (Phocoenoid)

hocoena).
 A;Reference number: A90415; MUID: 79000346; PMID: 687594
 A;Accession: A0415
 A;Molecule type: protein
 A;Residues: 56-90 <MEU>
 C;Superfamily: globin homology
 C;Keywords: chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier
 F;64/Domain: globin homology <GLB>
 F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted
 F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted
 Query Match 92.1%; Score 731; DB 1; Length 153;
 Best Local Similarity 90.8%; Pred. No. 1.4e-5%; Indels 0; Gaps 0;
 Matches 138; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
 Qy 2 LSEGEWQLVLHWAKEADVAGHQDILIRLFKSHPETLEKDFRKHKTAEAKMASEDL 61
 Db 2 LSEGEWQLVLNNGKVEDLAGHQDILIRLFKSHPETLEKDFRKHKTAEAKMASEDL 61
 Qy 62 KKHGTVTTLAIGAILKKGHHEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPG 121
 Db 62 KKHGTVTTLAIGAILKKGHHEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPA 121
 A;Reference number: A90415; MUID: 79000346; PMID: 687594
 A;Accession: B90415
 A;Molecule type: protein
 A;Species: Phocoenoides dalli (Dall's porpoise)
 C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 03-Mar-2000
 C;Accession: B90415; A02498
 R;Meuth, J.L.; Jones, B.N.; Garner, W.H.; Gurd, F.R.N.
 Biochemistry 17, 3429-3431, 1978
 A;Title: Complete amino acid sequence of the myoglobin from the Dall porpoise (Phocoenoid)
 A;Reference number: A90415; MUID: 79000346; PMID: 687594
 A;Accession: B90415
 A;Molecule type: protein
 A;Residues: 1-153 <MEU>
 C;Superfamily: globin homology
 C;Keywords: chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier
 F;2-147/Domain: globin homology <GLB>
 F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted
 F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted
 Query Match 92.1%; Score 731; DB 1; Length 153;
 Best Local Similarity 90.8%; Pred. No. 1.4e-5%; Indels 0; Gaps 0;
 Matches 138; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
 Qy 2 LSEGEWQLVLHWAKEADVAGHQDILIRLFKSHPETLEKDFRKHKTAEAKMASEDL 61
 Db 2 LSEGEWQLVLNNGKVEDLAGHQDILIRLFKSHPETLEKDFRKHKTAEAKMASEDL 61
 Qy 62 KKHGTVTTLAIGAILKKGHHEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPG 121
 Db 62 KKHGTVTTLAIGAILKKGHHEALKPLAQSHATKHKPIKYLEFISEAIHVVLHSRHPA 121
 A;Reference number: A90415; MUID: 79000346; PMID: 687594
 A;Accession: B90415
 A;Molecule type: protein
 A;Species: Balaenoptera acutorostrata (minke whale, lesser rorqual)
 C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 03-Mar-2000
 C;Accession: A02561
 R;Lehman, L.D.; Dwulet, F.E.; Bogardt Jr., R.A.; Jones, B.N.; Gurd, F.R.N.
 Biochemistry 16, 706-709, 1977

A;Title: The complete amino acid sequence of the major component myoglobin from the arch
A;Reference number: A02501 ; MUID:7712438 ; PMID:8368100
A;Molecule type: protein
A;Residues: 1-153 <LEH>
C;Comment: This myoglobin was isolated from skeletal muscle.
C;Superfamily: Globin; globin homology
C;Keywords: Chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier
F;2-147/Domain: Globin homology <GLB>
F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted
F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 92.1%; Score 731; DB 1; Length 153;
Best Local Similarity 90.8%; Pred. No. 1.4e-53;
Matches 139; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Qy	1 VLSGEQWLVHWWAKVEADVAGHGDILRLFKSHPTLEKFDKFHLKTEAEMKASED 60	Db	1 VLSDAEWHLVNLNWKVEADVAGHGDILRLFKGHPTLEKFDKFHLKTEAEMKASED 60
Qy	61 LKHGHTVLTALGAIKKKGHEAELKPLAQSHATKHKIPKYLEFISAEIIVLHSRHP 120	Db	61 LKHGHTVLTALGAIKKKGHEAELKPLAQSHATKHKIPKYLEFISAEIIVLHSRHP 120
Qy	121 GDFGADAQGMNKALELFRKDIAAKYKELGYQG 153	Db	121 AFGADAQAMNKALELFRKDIAAKYKELGFQG 153

RESULT 13
MYWHF
myoglobin - finback whale
C;Species: *Balaenoptera physalus* (finback whale, common rorqual)
C;Accession: A02502
R;Dimarchi, R.D.; Wang, C.C.; Hemmenway, J.B.; Gurd, F.R.N.
Biochemistry 17, 1968-1970, 1978
A;Title: Complete amino acid sequence of the major component myoglobin of finback whale
A;Reference number: A02502 ; MUID:78187212; PMID:656375
A;Molecule type: protein
A;Residues: 1-153 <DIM>
C;Comment: This myoglobin was isolated from skeletal muscle.
C;Superfamily: Globin; globin homology
C;Keywords: Chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier
F;2-147/Domain: Globin homology <GLB>
F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted
F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 91.4%; Score 726; DB 1; Length 153;
Best Local Similarity 90.4%; Pred. No. 3.5e-53;
Matches 138; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy	1 VLSGEQWLVHWWAKVEADVAGHGDILRLFKSHPTLEKFDKFHLKTEAEMKASED 60	Db	1 VLTDAEWHLVNLNWKVEADVAGHGDILSLFKGHPTLEKFDKFHLKTEAEMKASED 60
Qy	61 LKHGHTVLTALGAIKKKGHEAELKPLAQSHATKHKIPKYLEFISAEIIVLHSRHP 120	Db	61 LKHGHTVLTALGGLIKKKGHEAELKPLAQSHATKHKIPKYLEFISDAIIIVLHSRHP 120
Qy	121 GDFGADAQGMNKALELFRKDIAAKYKELGYQG 153	Db	121 ADFGADAQAMNKALELFRKDIAAKYKELGFQG 153

RESULT 14
MYWHU
myoglobin - Hubbs' whale
C;Species: Mesoplodon carlhubbsi ('Hubbs' whale)
C;Date: 27-Nov-1985 #sequence_revision 27-Nov-1985 #text_change 03-Mar-2000
C;Accession: A02505
R;Dwulet, J.A.; Dwalet, F.E.; Gurd, F.R.N.

Search completed: August 10, 2004, 15:29:27
Job time : 7.66828 secs

Biochim. Biophys. Acta 624, 121-129, 1980
A;Title: Complete amino acid sequence of the major component myoglobin from Hubb's beake
A;Reference number: A02505 ; MUID:81000592 ; PMID:7407230
A;Accession: A02505
A;Molecule type: protein
A;Residues: 1-153 <DWU>
C;Comment: This myoglobin was isolated from skeletal muscle.
C;Superfamily: Globin; globin homology
C;Keywords: Chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier
F;2-147/Domain: Globin homology <GLB>
F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted
F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 91.3%; Score 725; DB 1; Length 153;
Best Local Similarity 90.1%; Pred. No. 4.3e-53;
Matches 137; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

Qy	2 LSEGEGWQLVHWWAKVEADVAGHGDILRLFKSHPTLEKFDKFHLKTEAEMKASED 61	Db	2 LSEAWQVQLVHWWAKVEADLSGHGEELRLFKGHPTLEKFDKFHLKTEAEMKASED 61
Qy	62 KKHGHTVLTALGAIKKKGHEAELKPLAQSHATKHKIPKYLEFISAEIIVLHSRHP 121	Db	62 KKHGHTVLTALGGLIKKKGHEAELKPLAQSHATKHKIPKYLEFISDAIIIVLHSRHP 121
Qy	122 DFGADAGGMNKALELFRKDIAAKYKELGYQG 153	Db	122 DFGADAGGMNKALELFRKDIAAKYKELGFHQG 153

RESULT 15
MYWHZ
myoglobin - goose-beaked whale
C;Species: *Ziphius cavirostris* (goose-beaked whale)
C;Accession: A02504
R;Lehman, L.D.; Jones, B.N.; Dwulet, F.E.; Bogardt Jr., R.A.; Gurd, F.R.N.
Biochim. Biophys. Acta 625, 221-229, 1980
A;Title: Complete amino acid sequence of the major component myoglobin from the goose-beaked whale
A;Reference number: A02504 ; MUID:81063220; PMID:7437458
A;Accession: A02504
A;Molecule type: protein
A;Residues: 1-153 <LEH>
C;Comment: This myoglobin was isolated from skeletal muscle.
C;Superfamily: Globin; globin homology
C;Keywords: Chromoprotein; heme; iron; metalloprotein; muscle; oxygen carrier
F;2-147/Domain: Globin homology <GLB>
F;64/Binding site: oxygen (His) (distal axial ligand) #status predicted
F;93/Binding site: heme iron (His) (proximal axial ligand) #status predicted

Query Match 90.9%; Score 722; DB 1; Length 153;
Best Local Similarity 90.1%; Pred. No. 7.6e-53;
Matches 137; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Qy	2 LSEGEGWQLVHWWAKVEADVAGHGDILRLFKSHPTLEKFDKFHLKTEAEMKASED 61	Db	2 LSEAWQVQLVHWWAKVEADLSGHGEELRLFKGHPTLEKFDKFHLKTEAEMKASED 61
Qy	62 KKHGHTVLTALGAIKKKGHEAELKPLAQSHATKHKIPKYLEFISAEIIVLHSRHP 121	Db	62 KKHGHTVLTALGGLIKKKGHEAELKPLAQSHATKHKIPKYLEFISDAIIIVLHSRHP 121
Qy	122 DFGADAGGMNKALELFRKDIAAKYKELGYQG 153	Db	122 DFGADAGGMNKALELFRKDIAAKYKELGFHQG 153

B
I
G
A
N
T

Result No.	Score	Query Match Length	DB ID	Description
1	794	100.0	153	P02185 phyceratops caudatus muscle;
2	773	97.4	153	P02184 kogia simus muscle;
3	746	94.0	153	P02177 eschrichtius abrama muscle;
4	740	93.2	153	P02178 megaptera novaeangliae muscle;
5	738	92.9	153	P02181 inia geoffrensis muscle;
6	737	92.8	153	P02174 globicephala macrorhyncha muscle;
7	734	92.4	153	P02173 orcinus orca muscle;
8	733	92.3	153	P02172 tursiops truncatus muscle;
9	731	92.1	153	P02179 balaenoptera physalus muscle;
10	731	92.1	153	P02176 phocoenoides phocoena muscle;
11	726	91.4	153	P02180 balaenoptera physalus muscle;
12	725	91.3	153	P02183 mesoplodon europaeus muscle;
13	722	90.9	153	P14336 castor fiber muscle;
14	722	90.9	153	P02186 ziphodus caviae muscle;
15	717	90.3	153	P02188 equus caballus muscle;
16	707	89.0	153	P02145 pan troglodytes muscle;
17	705	88.8	153	P02166 periodicus hominis muscle;
18	703	88.5	153	P02170 orctocephalus macrourus muscle;
19	701	88.3	153	P02147 gorilla gorilla muscle;
20	701	88.3	153	P32456 tupaia glis muscle;
21	700	88.2	153	P02189 ondatra trivirgata muscle;
22	699	88.0	153	P02144 sus scrofa muscle;
23	695	87.5	153	P02150 macaca fasciata muscle;
24	694	87.4	153	P02148 pongo pygmaeus muscle;
25	693	87.3	153	P02146 hylobates agilis muscle;
26	693	87.3	153	P02165 tupaia glis muscle;
27	691	87.0	153	P02151 aotus trivirgatus muscle;
28	691	87.0	153	P02149 papio anubis muscle;
29	690	86.9	153	P02152 callithrix jacchus muscle;
30	690	86.9	153	P02153 roussetta roussetti muscle;
31	688	86.6	153	P02159 lemur lemur muscle;
32	688	86.6	153	P11343 lutreola lutreola muscle;
33	687	86.5	153	P02164 oryzceropus oryzceropus muscle;

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%
Listing First 45 summaries

Database : SwissProt_42.1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

#	Query	Match Length	DB ID	Description
1	VLSEGENQVLVHWAKEYAD.....ADELFRIKDIAXKYKEIUGYQG	153	P02167 nycticebus calyptronotus	P02161 zalophus californianus
2	MYG_PHYCA	153	P02164 MYG_ZALCA	P02165 CTIGU
3	ID MYG_PHYCA	153	P02163 MYG_HAIGR	P02162 halichoerus
4	AC P02155;	153	P02160 MYG_DELLE	P02168 delphinapterus
5	DT 21-JUL-1986 (Rel. 01, Created)	153	P02155 MYG_GALCR	P02153 galago crassicaudatus
6	DT 21-JUL-1986 (Rel. 01, Last sequence update)	153	P02153 MYG_CEDAP	P02154 apel
7	DT 15-MAR-2004 (Rel. 43, Last annotation update)	153	P02153 MYG_LAGLA	P02154 lagotrix lagotrichia
8	DB Myoglobin.	153	P02153 MYG_OCPFR	P04249 spalax leucodon
9	GN MB.	153	P02153 MYG_SPLBH	P02191 cervus elaphus
10	OS Phyceratops (Sperm whale) (Physeter macrocephalus).	153	P02153 MYG_CPRFL	P04249 prochimys
11	OC Bukarota; Metzcoa; Chordata; Craniata; Vertebrata; Butteleostomi; Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Physeteridae; Physeter.	153	P02153 MYG_PRGOU	

ALIGNMENTS

RESULT 1	MYG_PHYCA	STANDARD;	PRT;	153 AA.
	ID MYG_PHYCA			
	AC P02155;			
	DT 21-JUL-1986 (Rel. 01, Created)			
	DT 21-JUL-1986 (Rel. 01, Last sequence update)			
	DT 15-MAR-2004 (Rel. 43, Last annotation update)			
	DB Myoglobin.			
	GN MB.			
	OS Phyceratops (Sperm whale) (Physeter macrocephalus).			
	OC Bukarota; Metzcoa; Chordata; Craniata; Vertebrata; Butteleostomi; Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Physeteridae; Physeter.			
	NCBI_TaxID=9755;			
	RN [1] _			
	RP SEQUENCE.			
	RC TISSUE=Heart muscle;			
	RA Romero-Herrera A.E.; Lehmann H.;			
	RT "Residue 122 of sperm whale and horse myoglobin.";			
	RL Biochim. Biophys. Acta 336:318-323(1974).			
	RN [2]			
	RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF METMYOGLOBIN.			
	RX MEDLINE=77144097; PubMed=8455959;			
	RA Takano T.			
	RT "Structure of myoglobin refined at 2.0-A resolution. I. Crystallographic refinement of metmyoglobin from sperm whale."			
	RT deoxymyoglobin from sperm whale.";			
	RT Crystallographic refinement of metmyoglobin from sperm whale."			
	RL J. Mol. Biol. 110:537-568(1977).			
	RN [4]			
	RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF DEOXYMYOGLOBIN.			
	RX MEDLINE=77144098; PubMed=845560;			
	RA Takano T.			
	RT "Structure of myoglobin refined at 2.0-A resolution. II. Structure of deoxymyoglobin from sperm whale."			
	RT Crystallographic refinement of deoxymyoglobin from sperm whale."			
	RL J. Mol. Biol. 110:569-584(1977).			
	RN [5]			
	RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF OXYMYOGLOBIN.			
	RX MEDLINE=81119812; PubMed=7463482;			
	RA Lionetti C., Guanziroli M.G., Frigerio F., Ascenzi P., Bolognesi M.; Phillips S.E.V.			
	RT "X-ray crystal structure of the ferric sperm whale myoglobin: imidazole complex at 2.0-A resolution."			
	RL J. Mol. Biol. 142:331-354(1980).			
	RN [6]			
	RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).			
	RX MEDLINE=11934031;			
	RA Lionetti C., Guanziroli M.G., Frigerio F., Ascenzi P., Bolognesi M.; Phillips S.E.V.			
	RT "X-ray crystal structure of the ferric sperm whale myoglobin: imidazole complex at 2.0-A resolution."			
	RL J. Mol. Biol. 217:409-412(1991).			
	RN [7]			
	RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).			
	RX MEDLINE=99158603; PubMed=10049310;			

RA	Brunori M., Cutruzzola F., Savino C., Travaglini-Allocatelli C., Vallone B., Gibson Q.H.; "Structural dynamics of ligand diffusion in the protein matrix: a study on a new myoglobin mutant Y(B10) Q(E10)." ; R(B10)."; [8]	
RL		
RN		
RP	X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).	
RX	MEDLINE=20160904; PubMed=10681426;	
RA	Brunori M., Vallone B., Cutruzzola F., Travaglini-Allocatelli C., Berendzen J., Chu K., Sweet R.M., Schlichting I.; "The role of cavities in protein dynamics: crystal structure of a photolytic intermediate of a mutant myoglobin." ; Proc. Natl. Acad. Sci. U.S.A. 97:2058-2063 (2000).	
RT		
RL	-1- FUNCTION: Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.	
CC	-1- SIMILARITY: Belongs to the globin family.	
CC	PIR; A90591; MYHBP.	
DR	PDB; 101PM; 08-APR-98.	
DR	PDB; 102M; 08-APR-98.	
DR	PDB; 103M; 08-APR-98.	
DR	PDB; 104M; 08-APR-98.	
DR	PDB; 105M; 08-APR-98.	
DR	PDB; 106M; 08-APR-98.	
DR	PDB; 107M; 08-APR-98.	
DR	PDB; 108M; 20-MAY-98.	
DR	PDB; 109M; 08-APR-98.	
DR	PDB; 110M; 08-APR-98.	
DR	PDB; 111M; 08-APR-98.	
DR	PDB; 112M; 08-APR-98.	
DR	PDB; 1A6G; 21-OCT-98.	
DR	PDB; 1A6K; 06-APR-99.	
DR	PDB; 1A6M; 06-APR-99.	
DR	PDB; 1A6N; 06-APR-99.	
DR	PDB; 1ABS; 01-APR-97.	
DR	PDB; 1AJG; 12-NOV-97.	
DR	PDB; 1AJH; 12-NOV-97.	
DR	PDB; 1BVC; 31-JUL-95.	
DR	PDB; 1BVD; 31-JUL-95.	
DR	PDB; 1BZ6; 11-NOV-98.	
DR	PDB; 1BZP; 10-MAY-99.	
DR	PDB; 1BZR; 10-MAY-99.	
DR	PDB; 1CH1; 09-APR-99.	
DR	PDB; 1CH2; 09-APR-99.	
DR	PDB; 1CH3; 09-APR-99.	
DR	PDB; 1CH5; 09-APR-99.	
DR	PDB; 1CH7; 09-APR-99.	
DR	PDB; 1CH9; 09-APR-99.	
DR	PDB; 1C1K; 09-APR-99.	
DR	PDB; 1C1O; 09-APR-99.	
DR	PDB; 1C08; 14-JUN-99.	
DR	PDB; 1C09; 14-JUN-99.	
DR	PDB; 1CPO; 14-JUN-99.	
DR	PDB; 1CP5; 14-JUN-99.	
DR	PDB; 1CPW; 14-JUN-99.	
DR	PDB; 1CQ2; 16-AUG-99.	
DR	PDB; 1D01; 02-APR-00.	
DR	PDB; 1D03; 02-APR-00.	
DR	PDB; 1D04; 02-APR-00.	
DR	PDB; 1DD7; 02-APR-00.	
DR	PDB; 1DTI; 30-JAN-00.	
DR	PDB; 1DTM; 14-JUN-00.	
DR	PDB; 1DUK; 08-APR-03.	
DR	PDB; 1DUD; 27-DEC-00.	
DR	PDB; 1IXC; 02-APR-00.	
DR	PDB; 1JDD; 20-JUL-00.	
DR	PDB; 1JBC; 13-AUG-99.	
DR	PDB; 1J63; 19-JUL-00.	
DR	PDB; 1J6H; 21-JUL-00.	
DR	PDB; 1JCS; 31-OCT-93.	
DR	PDB; 1JUT; 12-NOV-97.	
DR	PDB; 1IOP; 08-APR-98.	
DR	PDB; 1IRC; 11-JUL-96.	
PDB	1JDO; 27-MAY-98.	
PDB	1JP6; 16-JAN-02.	
PDB	1JPB; 16-JAN-02.	
PDB	1JP9; 16-JAN-02.	
PDB	1JPB; 16-JAN-02.	
PDB	1JWB; 10-OCT-01.	
PDB	1L2K; 21-AUG-02.	
PDB	1LTW; 23-DEC-96.	
PDB	1MBC; 09-JAN-89.	
PDB	1MBD; 27-OCT-83.	
PDB	1MBI; 15-OCT-91.	
PDB	1MBN; 27-OCT-83.	
PDB	1MBO; 30-SEP-83.	
PDB	1MCY; 07-DEC-95.	
PDB	1MGN; 31-JAN-94.	
PDB	1MLF; 31-AUG-94.	
PDB	1MLG; 31-AUG-94.	
PDB	1MLH; 31-AUG-94.	
PDB	1MLJ; 31-AUG-94.	
PDB	1MLK; 31-AUG-94.	
PDB	1MLL; 31-AUG-94.	
PDB	1MLM; 31-AUG-94.	
PDB	1MLN; 31-AUG-94.	
PDB	1MLO; 31-AUG-94.	
PDB	1MLQ; 07-FEB-95.	
PDB	1MLR; 31-AUG-94.	
PDB	1MLS; 31-AUG-94.	
PDB	1MLU; 31-AUG-94.	
PDB	1MOA; 07-FEB-95.	
PDB	1MOP; 07-FEB-95.	
PDB	1MOC; 07-FEB-95.	
PDB	1MOD; 07-FEB-95.	
PDB	1MOT; 07-FEB-95.	
PDB	1MTI; 15-SEP-95.	
PDB	1MTJ; 15-SEP-95.	
PDB	1MTK; 15-SEP-95.	
PDB	1MTP; 27-FEB-95.	
PDB	1MWF; 27-FEB-95.	
PDB	1MYM; 31-JAN-94.	
PDB	1OBM; 08-APR-98.	
PDB	1OIJ; 27-MAY-98.	
PDB	1OFP; 11-NOV-98.	
PDB	1ISPE; 08-MAR-96.	
PDB	1SNM; 31-JAN-94.	
PDB	1TWS; 08-NOV-96.	
PDB	1VXA; 01-AUG-96.	
PDB	1VXB; 01-AUG-96.	
PDB	1VXH; 01-AUG-96.	
PDB	1YOG; 01-AUG-96.	
PDB	1VXD; 01-AUG-96.	
PDB	1VXE; 01-AUG-96.	
PDB	1VXF; 01-AUG-96.	
PDB	1ZMB5; 15-APR-91.	
PDB	2MBW; 23-DEC-96.	
PDB	2MGAA; 31-JAN-94.	
PDB	2MGH; 31-JUL-94.	
PDB	2MGI; 31-JUL-94.	
PDB	2MGJ; 31-JUL-94.	
PDB	2MGK; 31-JUL-94.	
PDB	2MGL; 31-JUL-94.	
PDB	2MGM; 31-JUL-94.	
PDB	2NYA; 31-JAN-94.	
PDB	2NYB; 31-JAN-94.	
PDB	2NYC; 31-JAN-94.	
PDB	2NYD; 31-JAN-94.	
PDB	2NYE; 31-JAN-94.	
PDB	2NYF; 31-JAN-94.	
PDB	2NYG; 31-JAN-94.	
PDB	2NYH; 31-JAN-94.	
PDB	2NYI; 07-DEC-96.	
PDB	2NYK; 01-DEC-96.	
PDB	2NYL; 07-DEC-96.	
PDB	2NYM; 31-JUL-94.	
PDB	2NYN; 31-JUL-94.	
PDB	2NYO; 31-JUL-94.	
PDB	2NYP; 31-JUL-94.	
PDB	2NYQ; 31-JUL-94.	
PDB	2NYR; 31-JUL-94.	
PDB	2NYT; 31-JUL-94.	
PDB	2NYU; 31-JUL-94.	
PDB	2NYV; 31-JUL-94.	
PDB	2NYW; 31-JUL-94.	
PDB	2NYX; 31-JUL-94.	
PDB	2NYZ; 31-JUL-94.	

Query Match		Score 794; DB 1; Length 153;	
Best Local Similarity 100.0%; Pred. No. 1.e-61;		Indels 0; Gaps 0;	
Matches 153; Conservative 0; Mismatches 0;		De Myoglobin.	
Qy	1 VLSGEWQVLVHWWAKVEADYAGHGODILIRLFLKSFETLEKFDRLKHLKTEAEMKASED 60	DB	1 VLSGEWQVLVHWWAKVEADYAGHGODILIRLFLKSFETLEKFDRLKHLKTEAEMKASED 60
Qy	1 VLSGEWQVLVHWWAKVEADYAGHGODILIRLFLKSFETLEKFDRLKHLKTEAEMKASED 60	Db	1 VLSGEWQVLVHWWAKVEADYAGHGODILIRLFLKSFETLEKFDRLKHLKTEAEMKASED 60
Qy	61 LKKHGTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120	Db	61 LKKHGTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120
Qy	61 LKKHGTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120	Db	61 LKKHGTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120
RESULT 2			
AC	MGKOGSI	MYG KOGSI	STANDARD;
AC	P02184;	PRT;	153 AA.
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DE	10-OCT-2003 (Rel. 42, Last annotation update)		
MB			
OS	Kogia simus (Dwarf sperm whale).		
OC	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Buteleostomi;		
OC	Physeteridae; Kogia.		
NCBI_TaxID	=9752;		
[1]			
SEQUENCE,			
RA	MEMLINE=771134684; PubMed=843520;		
RA	Dwullet F.E., Jones B.N., Lehman I.D., Gurd F.R.N.;		
RT	"The complete amino acid sequence of the major component myoglobin of dwarf sperm whale (Kogia simus).";		
RT	Biochemistry 16:873-877(1977)		
CC	-!- FUNCTION: Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.		
CC	-!- SIMILARITY: Belongs to the globin family.		
PIR; A90407; MYHWH.			
InterPro; IPR000971; Globin.			
InterPro; IPR002335; Myoglobin.			
Pfam; PF0042; globin; 1.			
PRINTS; PRO0613; MYOGLOBIN.			
PROSITE; PS01033; GLOBIN; 1.			
KW	Heme; Oxygen transport; Transport; Muscle		
FT	METAL 64		IRON (HEME DISTAL LIGAND).
FT	METAL 64		IRON (HEME PROXIMAL LIGAND).
SQ	SEQUENCE 153 AA; 17237 MW;		4A5587BAE05C7D CRC64;
Query Match 94.0%; Score 746; DB 1; Length 153;			
Best Local Similarity 92.8%; Pred. No. 1.e-57;	Matches 142; Conservative 6; Mismatches 5; Indels 0; Gaps 0;		
Matches 142; Conservative			
Qy	1 VLSEGEWQVLVHWWAKVEADYAGHGODILIRLFLKSFETLEKFDRLKHLKTEAEMKASED 60	DB	1 VLSEDAEWQVLVHWWAKVEADYAGHGODILIRLFLKSFETLEKFDRLKHLKTEAEMKASED 60
Qy	61 LKKHGTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120	DB	61 LKKHGNTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120
Qy	61 LKKHGNTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120	DB	61 LKKHGNTVTTLGALKLKKKGHEAEKLPLAQSHATKHKIPKYLEFSEATHVLHSRHP 120
Qy	121 GDFGADAQGMNKALELFRKDIAKYKELGYQ 153	DB	121 GDFGADAQGMNKALELFRKDIAKYKELGYQ 153
RESULT 4			
MYG MEGNO	MYG MEGNO	STANDARD;	PRT;
ID	AC P02178;		153 AA.
RT	21-JUL-1986 (Rel. 01, Created)		
RT	21-JUL-1986 (Rel. 01, Last sequence update)		
DE	16-OCT-2001 (Rel. 40, Last annotation update)		
GN	Myoglobin.		
OS	Megaptera novaengliae (Humpback whale).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;		
OC	Balaenopteridae; Megapteridae.		
NCBI_TaxID	=9773;		
[1]			
RN	SEQUENCE.		
RC	TISSUE=Skeletal muscle;		
RX	Medline=77901601; PubMed=698193;		
RA	Lehman L.D., Dwullet F.E., Jones B.N., Bogardt R.A. Jr., Krueckeberg S.T., Visscher R.B., Gurd F.R.N.;		
RA	"Complete amino acid sequence of the major component myoglobin from the humpback whale. Meagan era		
RT	"novoangliae"		
RESULT 3			
MYG ESCGI	MYG ESCGI	STANDARD;	PRT;
ID	121 ADFGADAQGMNKALELFRKDIAKYKELGYQ 153		153 AA.
DB	121 ADFGADAQGMNKALELFRKDIAKYKELGYQ 153		

RL	Biochemistry 17:3736-3739(1997).	KW	Heme; Oxygen transport; Transport; Muscle.
RN	[2]	FT	IRON (HEME DISTAL LIGAND).
RP	SEQUENCE OF 1-60.	FT	METAL 64 64 IRON (HEME PROXIMAL LIGAND).
RX	PUBMED=6059350;	FT	METAL 93 93 IRON (HEME PROXIMAL LIGAND).
RA	Edman P.; Begg G.;	SQ	SEQUENCE 153 AA; 17071 MW; 7222ED4EBCB5D CRC64;
RT	"A protein sequenator.";	Query Match	92.9%; Score 738; DB 1; Length 153;
RL	Bur. J. Biochem. 1:80-91(1967).	Best Local Similarity	91.4%; Pred. No. 7.7e-57;
CC	- - FUNCTION: Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.	Matches	139; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
CC	- - MISCELLANEOUS: This sequence was the first determined using the Edman sequencing system.	Qy	2 LSEGEGNVLVHWAKEYADVAHGQDILIRLFLKSHPETLEKFDRFKHLKTEAEMKASEDL 61
CC	- - SIMILARITY: Belongs to the globin family.	Db	2 LSDGEWOLVNWGYEADLGHGQDILIRLFLKGHPETLEKFDRFKHLKTEAEMKASEDL 61
DR	PIR; A90416; MYWHH.	Qy	62 KKHGHTVLTAGAILKKKGHHAEELPLAQSHATKHKIPIKYLEFISEAIIHVLSRHPG 121
DR	HSSP; P02145; 1BZ6.	Db	62 RKGHNTVLTAGIILRKKGHHAEELPLAQSHATKHKIPIKYLEFISEAIIHVLSRHPG 121
DR	InterPro; IPR000971; Globin.	Qy	122 DFGADAOGMANNKALELFRKDIAAKYKELGYQG 153
DR	PFam; PF00042; globin; 1.	Db	122 DFGADAOGMANNKALELFRKDIAAKYKELGFHQG 153
DR	PRINTS; PR00613; MYOGLOBIN.	Qy	
DR	PROSITE; PS01033; GLOBIN; 1.	Db	
DR	DR; Heme; Oxygen transport; Transport; Muscle.	RESULT 6	
FT	METAL 64 64 IRON (HEME DISTAL LIGAND).	MYG_GLOME	
FT	METAL 93 93 IRON (HEME PROXIMAL LIGAND).	ID	MYG_GLOME
SQ	SEQUENCE 153 AA; 17132 MW; 4FAFB81AOE805CTD CRC64;	ID	STANDARD;
Query Match	93.2%; Score 740; DB 1; Length 153;	PRT	153 AA.
Best Local Similarity	92.2%; Pred. No. 5.2e-37;	AC	P02174;
Matches	141; Conservative 6; Mismatches 6; Indels 0; Gaps 0;	DT	21-JUL-1986 (Rel. 01, Created)
Qy	VLSSEGEGNVLVHWAKEYADVAHGQDILIRLFLKSHPETLEKFDRFKHLKTEAEMKASEDL 60	DT	21-JUL-1986 (Rel. 01, Last sequence update)
Db	VLSDAEFLQVLNWIWGYEADVAHGQDILIRLFLKGHPETLEKFDRFKHLKTEAEMKASEDL 60	DT	16-OCT-2001 (Rel. 40, Last annotation update)
DR	IKKHGTHTVLTAGAILKKKGHHAEELPLAQSHATKHKIPIKYLEFISEAIIHVLSRHP 120	DE	Myoglobin.
DR	LKGHNTVLTAGIILRKKGHHAEELPLAQSHATKHKIPIKYLEFISEAIIHVLSRHP 120	GN	GB
DR	Globicephala melaena (Long-finned pilot whale) (Globicephala melas).	OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buteraria; Cetacea; Odontoceti; Delphinidae; Globicephala.
RN	[1]	OC	
RN	NCBI_TaxID=9731;	OC	
RN	SEQUENCE.	OX	
RP	RC TISSUE=Skeletal muscle;	OX	
RC	MEDLINE=78187213; PubMed=656376;	RX	
RC	Jones B.N., Dwulet F.E., Lehman L.D., Garner M.H., Bogardt R.A. Jr., Garner W.H., Gurd F.R.N.; "Complete amino acid sequence of myoglobin from the pilot whale, Globicephala melaena"; Biochemistry 17:1971-1974 (1978).	RA	
RC	- - FUNCTION: Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.	RA	
CC	- - SIMILARITY: Belongs to the globin family.	RT	
CC	DR; P02496; MYWT.	RI	
CC	DR; InterPro; IPR000971; Globin.	DR	
CC	DR; Pfam; PF00042; globin; 1.	DR	
CC	DR; PRINTS; PR00613; MYOGLOBIN.	DR	
CC	DR; PROSITE; PS01033; GLOBIN; 1.	DR	
CC	DR; Heme; Oxygen transport; Transport; Muscle.	DR	
CC	FT METAL 64 64 IRON (HEME DISTAL LIGAND).	FT	
CC	FT METAL 93 93 IRON (HEME PROXIMAL LIGAND).	FT	
CC	FT SEQUENCE 153 AA; 17085 MW; 510980DECB83DFD CRC64;	FT	
RA	Dwulet F.B., Bogardt R.A., Jones B.N., Lehman L.D., Gurd F.R.N.; "The complete amino acid sequence of the major component myoglobin of Amazon river dolphin (<i>Inia geoffrensis</i>)."; Biochemistry 14:5336-5343 (1975).	RA	
RT	- - FUNCTION: Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.	RT	
CC	- - SIMILARITY: Belongs to the globin family.	CC	
CC	DR; P02503; MYDAR.	DR	
CC	DR; InterPro; IPR000971; Globin.	DR	
CC	DR; Pfam; PF00042; globin; 1.	DR	
CC	DR; PRINTS; PR00613; MYOGLOBIN.	DR	
CC	DR; PROSITE; PS01033; GLOBIN; 1.	DR	
Qy	2 LSEGEGNVLVHWAKEYADVAHGQDILIRLFLKSHPETLEKFDRFKHLKTEAEMKASEDL 61	Qy	
Db	2 LSDGEWOLVNWGYEADLGHGQDILIRLFLKGHPETLEKFDRFKHLKTEAEMKASEDL 61	Db	
Qy	62 KKHGHTVLTAGAILKKKGHHAEELPLAQSHATKHKIPIKYLEFISEAIIHVLSRHPG 121	Qy	
Db	62 KKGHNTVLTAGIILRKKGHHAEELPLAQSHATKHKIPIKYLEFISEAIIHVLSRHPG 121	Db	
Qy	122 DFGADAOGMANNKALELFRKDIAAKYKELGYQG 153	Qy	

Db	122	EFFGADAGQAMKAKALEFRKDIAAKYKELGFGH	153	
RESULT 7				
MYG_ORCOR	STANDARD;	PRT;	153 AA.	
ID	MYG_ORCOR			
AC	P02173;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DE	Myoglobin.			
GN	MB.			
OC	Orcinus orca (Killer whale).			
OC	Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;			
OC	Orcinus.			
OX	NCBI_TaxID=9733;			
RN				
RP	SEQUENCE:			
RC	TISSUE=Skeletal muscle;			
RX	MEDLINE=81267443; PubMed=6113067;			
RA	Castillo O., Lehmann H., Jones L.T.;			
RT	"The myoglobin of the killer whale (Orcinus orca).";			
RL	Biochim. Biophys. Acta 491:23-28 (1977).			
CC	-!- FUNCTION: Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.			
CC	-!- SIMILARITY: Belongs to the globin family.			
DR	PIR; A92986; MYWHL.			
DR	HSSP; P02185; 1A6M.			
DR	InterPro; IPR000971; Globin.			
DR	PF0042; Globin.			
DR	PRINTS; PR00613; MYOGLOBIN.			
DR	PROSITE; PS01033; GLOBIN_1.			
KW	Oxygen transport; Transport; Muscle.			
FT	METAL 64 64 IRON (HEME DISTAL LIGAND).			
FT	METAL 93 93 IRON (HEME PROXIMAL LIGAND).			
FT	CONFLICT 122 122 E -> Q (IN REF. 2).			
SQ	SEQUENCE 153 AA; 17071 MW; A408C49CB83C4C CRC64;			
Query Match				
Qy	Best Local Similarity 92.4%; Score 734; DB 1; Length 153;			
Matches	139; Conservative 8; Mismatches 5; Indels 0; Gaps 0;			
Qy	2 LSEGEWQLVNHWAKEYADYAGHGODILIRLFLKSHBETLKFDRFKHLKTEADMKAESDL 61			
Db	2 KKHGNTVLTALGAIKKKGHEAEIKPLAQSHATHKHKIPKYLEFTSEAITHVLSRHPG 121			
Qy	62 DFGADQGAMKAKALEFRKDIAAKYKELGFGH 153			
Db	122 EFFGADAGQAMKAKALEFRKDIAAKYKELGFGH 153			
RESULT 8				
MYG_TURTR	STANDARD;	PRT;	153 AA.	
ID	MYG_TURTR			
AC	P02172; P02175;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			

DB	DE Myoglobin.	OS MB.
	Tursiops truncatus (Atlantic bottle-nosed dolphin), and Tursiops truncatus (Black sea dolphin).	OS Delphinus delphis (Saddleback dolphin) (Black sea dolphin).
	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;	OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
	OC Tursiops.	OC Tursiops.
	OX NCBI_TaxID=9733;	OX NCBI_TaxID=9739; 9728;
	RN	RN
	RP	RP
	SEQUENCE:	SEQUENCE=T_truncatus; TISSUE=Skeletal muscle;
	RC	RC SPECIES=D_delphinus; TISSUE=Skeletal muscle;
	RX	RX MEDLINE=8020893; PubMed=911808;
	RA	RA Wang C.-C., Avila R., Jones B.N., Gurd F.R.N.;
	RT	RT "Complete primary structure of the major component myoglobin of Pacific common dolphin (Delphinus delphis).";
	RL	RL Biochemistry 16:4978-4981(1977).
	RN	RN [3]
	RP	RP [2]
	SEQUENCE:	SEQUENCE OF 1-31.
	RC	RC SPECIES=D_delphinus;
	RX	RX Klub I., Bakardjieva A.;
	RA	RA "Primary structure of N-terminal part of molecule of dolphin myoglobin.";
	RT	RT FEBS Lett. 17:31-34 (1971).
	RN	RN [4]
	RP	RP SEQUENCE OF 31-153.
	RC	RC SPECIES=D_delphinus;
	RX	RX MEDLINE=70104229; PubMed=5473803;
	RA	RA Karadjova M., Nedkov P., Bakardjieva A.; Genov N.;
	RT	RT "Differences in amino acid sequence between dolphin and sperm whale myoglobins.";
	RL	RL Biolog. Biophys. Acta 221:136-139 (1970).
	RN	RN REVISION TO 26.
	RP	RP SPECIES=D_delphinus;
	RC	RC RA Klub I.;
	RX	RX Subhitt (DEC-1977) to the PIR data bank.
	RA	RA CC -!- FUNCTION: Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.
	CC	CC -!- SIMILARITY: Belongs to the globin family.
	DR	DR PIR; A02494; MYDD.
	DR	DR PIR; A26230; MYDDBS.
	DR	DR HSSP; P02185; 1B26.
	DR	DR InterPro; IPR000971; Globin.
	DR	DR InterPro; IPR002335; Myoglobin.
	DR	DR Pfam; PF00042; Globin; 1.
	DR	DR PRINTS; PR00613; MYOGLOBIN.
	DR	DR PROSITE; PS01033; GLOBIN_1.
	DR	DR PIR; A02494; MYDD.
	KW	KW Heme; Oxygen transport; Transport; Muscle.
	FT	FT METAL 64 64 IRON (HEME DISTAL LIGAND).
	FT	FT METAL 93 93 IRON (HEME PROXIMAL LIGAND).
	FT	FT CONFLICT 21 21 L -> V (IN REF. 3).
	FT	FT CONFLICT 28 28 V -> I (IN REF. 3).
	FT	FT CONFLICT 66 66 N -> D (IN REF. 4).
	FT	FT CONFLICT 122 122 E -> Q (IN REF. 4).
	SQ	SQ SEQUENCE 153 AA; 17057 MW; A408C4895BBAC53 CRC64;
		Query Match 92.3%; Score 733; DB 1; Length 153;
		Best Local Similarity 90.8%; Pred. No. 2.1e-56;
		Matches 138; Conservative 9; Mismatches 5; Indels 0; Gaps 0;
Qy	2 LSEGEWQLVNHWAKEYADYAGHGODILIRLFLKSHBETLKFDRFKHLKTEADMKAESDL 61	
Db	2 KKHGNTVLTALGAIKKKGHEAEIKPLAQSHATHKHKIPKYLEFTSEAITHVLSRHPG 121	
Qy	62 DFGADQGAMKAKALEFRKDIAAKYKELGFGH 153	
Db	122 EFFGADAGQAMKAKALEFRKDIAAKYKELGFGH 153	
	Qy	2 LSEGEWQLVNHWAKEYADYAGHGODILIRLFLKSHBETLKFDRFKHLKTEADMKAESDL 61
Db	2 LSDGEWQLVNHWAKEYADYAGHGODILIRLFLKSHBETLKFDRFKHLKTEADMKAESDL 61	

DR InterPro; IPR000971; Globin.
 DR InterPro; IPR002335; Myoglobin.
 DR Pfam; PF00042; Globin; 1.
 DR PRINTS; PRO0613; MYOGLOBIN; 1.
 DR PROSITE; PS01033; GLOBIN; 1.
 KW Heme; Oxygen transport; Transport; Muscle; 3D-structure.
 PT METAL 64 64 IRON (HEME DISTAL LIGAND).
 PT METAL 93 93 IRON (HEME PROXIMAL LIGAND).
 FT CONFLICT 122 122 D -> N (IN REF. 1).
 FT HELIX 4 18
 FT TURN 19 20
 FT HELIX 21 35
 FT HELIX 37 40
 FT TURN 41 42
 FT TURN 44 48
 FT HELIX 52 57
 FT HELIX 59 77
 FT TURN 78 80
 FT HELIX 83 94
 FT TURN 95 96
 FT HELIX 101 118
 FT TURN 120 121
 FT HELIX 125 148
 FT TURN 149 150
 SQ SEQUENCE 153 AA; 16951 MW; 89CA01974231E93C CRC64;
 Query Match 90.3%; Score 717; DB 1; Length 153;
 Best Local Similarity 88.2%; Pred. No. 5e-55;
 Matches 134; Conservative 9; Mismatches 9; Indels 0; Gaps 0;
 QY 2 LSEGENQVLVLYWAKVEADVAHGQDILIRLFLFKSHPBTILEFDREKHLKTEAEMKASEDL 61
 Db 2 LSDGEVQQVLNWKGVEADLGCHGQVLRUFTGAPBTILEFDKPFKHLKTEAEMKASEDL 61
 QY 62 KKHGTIVVLTAGAILKKGGHEAEELKPLAQSHATRKHPIPKYLEFTSEAIHVLSRHPG 121
 Db 62 KKHGTIVVLTAGGLKKGGHEAEELKPLAQSHATRKHPIPKYLEFTISDAI HVLSKHPG 121
 QY 122 DFGADAOGMNKAELFRKDIAAKYKELGYQG 153
 Db 122 DFGADAOGMNKAELFRNDIAAKYKELGFQG 153

Search completed: August 10, 2004, 15:26:50
 Job time : 5.8598 secs

B
l
a
n
k

Qy	66	TTVLTAAGLILKKKGHEAELKPLAQSHATKHKIPIKYLEFISEAIIHVLHSRHPGDFGA	125	Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Scombroidei;
Db	62	ATVLKKGELLKAKGHNDAILKPMANSHAATKHKIPINNFKLISEIIVKMQEAGMDA-	120	Scombridae; Thunnus.
Qy	126	DAQAMMNKAELFRKDIAAKYKELGYQG	153	NCBI_TaxID:8236;
Db	121	-GQQALRNVMMAVIADLEANYKELGFSG	147	RN
SEQUENCE FROM N.A.				
RESULT 5				
Q9DGJ7	ID	Q9DGJ7 ; PRELIMINARY;	PRT;	147 AA.
AC	DR	Q9DGJ7 ;	RC	TISSUE=Skeletal muscle;
DT	DR	01-MAR-2001 (TREMBLrel. 16, Created)	RA	RX MEDLINE=21146139; PubMed=11247835;
DT	DR	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	RA	RA Marcinek D.J.; Bonaventura J.; Wittenberg J.B.; Block B.A.;
DT	DR	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	RT	RT "Oxygen affinity and amino acid sequence of myoglobins from endothermic and ectothermic fish.";
DE	DR	Myoglobin.	RL	RL Am. J. Physiol. Regul. Integr. Comp. Physiol. 280:R1123-R1133 (2001).
OS	DR	Thunnus albacares (Yellowfin tuna) (Neothunnus macropterus).	DR	DR EMBL: AF291832; AAG021061; - .
OC	DR	Bukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi;	DR	HSSP: P02205; 1 MYT.
OC	DR	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	DR	GO; GO:0005344; P: oxygen transporter activity; IEA.
OC	DR	Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Scombroidei;	DR	GO; GO:0006810; P: oxygen transport; IEA.
OC	DR	Scombridae; Thunnus.	DR	InterPro; IPR00971; Globin.
ON	DR	[1] _ TaxID:8236;	DR	InteroPro; IPR02335; Myoglobin.
Qy	126	SEQUENCE FROM N.A.	DR	Pfam: PF00042; Globin_1.
Db	121	SEQUENCE FROM N.A.	DR	PRINTS: PR0613; MYOGLOBIN.
Qy	126	SEQUENCE FROM N.A.	DR	PROSITE; PS01033; GLOBIN_1.
Db	121	SEQUENCE FROM N.A.	DR	Heme; Oxygen transport; Transport.
Qy	126	SEQUENCE FROM N.A.	DR	KW Score: 147 AA; 15667 MW; 26114A4B946A1191 CRC64;
Db	121	SEQUENCE FROM N.A.	DR	SQ
Qy	126	SEQUENCE FROM N.A.	DR	Query Match 36.3%; Score 288.5; DB 13; Length 147;
Db	121	SEQUENCE FROM N.A.	DR	Best Local Similarity 41.2%; Pred. No. 4.4e-19;
Qy	126	SEQUENCE FROM N.A.	DR	Matches 61; Conservative 25; Mismatches 59; Indels 3; Gaps 2;
Db	121	SEQUENCE FROM N.A.	DR	Qy 6 EWQLVNHWKVYEADYAGHGDDILIRFLKSPETLEKFDRFKHLTKPAEMKASEDILKKHG 65
Qy	126	SEQUENCE FROM N.A.	DR	3 DEDAVLKCGWGEVEADYTIGUVLTFLKEPDTQKLPKFGI -AQADLAGNAASAHG 61
Db	121	SEQUENCE FROM N.A.	DR	66 VTVALTALGILKKKGHEAEKLPLAQSHATKHKIPKYLEFISEAIIHVLHSRHPGDFGA 125
Qy	126	SEQUENCE FROM N.A.	DR	62 ATVLKKGELLKAKGSHASILKPMASHATKHKIPNNFKLISEVLYKMQEAGLDAG- 120
Db	121	SEQUENCE FROM N.A.	DR	Qy 126 DAQAMMNKAELFRKDIAAKYKELGYQG 153
Qy	126	SEQUENCE FROM N.A.	DR	Db 121 -GQTALRNNGIIIADEANEKYKELGTG 147
RESULT 7				
Q9DGJ8	ID	Q9DGJ8 ; PRELIMINARY;	PRT;	146 AA.
AC	DR	Q9DGJ8 ;	RC	TISSUE=Skeletal muscle;
DT	DT	01-MAR-2001 (TREMBLrel. 16, Created)	RA	RX MEDLINE=21146139; PubMed=11247835;
Matches 65	DR	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	RA	RA Marcinek D.J.; Bonaventura J.; Wittenberg J.B.; Block B.A.;
Conservative 22	DR	01-MAR-2003 (TREMBLrel. 24, Last annotation update)	RT	RT "Oxygen affinity and amino acid sequence of myoglobins from endothermic and ectothermic fish.";
Mismatches 58	DR	01-MAR-2003 (TREMBLrel. 24, Last annotation update)	RL	RL Am. J. Physiol. Regul. Integr. Comp. Physiol. 280:R1123-R1133 (2001).
Indels 3	DR	01-MAR-2003 (TREMBLrel. 24, Last annotation update)	DR	DR EMBL: AF291837; AAG021111; - .
Gaps 2	DR	01-MAR-2003 (TREMBLrel. 24, Last annotation update)	DR	HSSP: P02205; 1 MYT.
Qy	126	SEQUENCE FROM N.A.	DR	GO; GO:0005344; P: oxygen transporter activity; IEA.
Db	121	SEQUENCE FROM N.A.	DR	GO; GO:0015671; P: oxygen transport; IEA.
Qy	126	SEQUENCE FROM N.A.	DR	InterPro; IPR002335; Myoglobin.
Db	121	SEQUENCE FROM N.A.	DR	Pfam: PF00042; Globin_1.
Qy	126	SEQUENCE FROM N.A.	DR	PRINTS: PR0613; MYOGLOBIN.
Db	121	SEQUENCE FROM N.A.	DR	PROSITE; PS01033; GLOBIN_1.
Qy	126	SEQUENCE FROM N.A.	DR	Heme; Oxygen transport; Transport.
Db	121	SEQUENCE FROM N.A.	DR	KW Score: 147 AA; 15660 MW; 3D13814403B48D44 CRC64;
Qy	126	SEQUENCE FROM N.A.	DR	SQ
Qy	126	SEQUENCE FROM N.A.	DR	Query Match 38.5%; Score 305.5; DB 13; Length 147;
Db	121	SEQUENCE FROM N.A.	DR	Best Local Similarity 43.9%; Pred. No. 1.1e-20;
Qy	126	SEQUENCE FROM N.A.	DR	Matches 65; Conservative 22; Mismatches 58; Indels 3; Gaps 2;
Db	121	SEQUENCE FROM N.A.	DR	Qy 6 EWQLVNHWKVYEADYAGHGDDILIRFLKSPETLEKFDRFKHLTKPAEMKASEDILKKHG 65
Qy	126	SEQUENCE FROM N.A.	DR	3 DEDAVLKCGWGEVEADYTIGUVLTFLKEPDTQKLPKFGI -AQADLAGNAASAHG 61
Db	121	SEQUENCE FROM N.A.	DR	66 VTVALTALGILKKKGHEAEKLPLAQSHATKHKIPKYLEFISEAIIHVLHSRHPGDFGA 125
Qy	126	SEQUENCE FROM N.A.	DR	62 ATVLKKGELLKAKGSHASILKPMASHATKHKIPNNFKLISEVLYKMQEAGLDAG- 120
Db	121	SEQUENCE FROM N.A.	DR	Qy 126 DAQAMMNKAELFRKDIAAKYKELGYQG 153
Qy	126	SEQUENCE FROM N.A.	DR	Db 121 -GQTALRNNGIIIADEANEKYKELGTG 147
RESULT 6				
Q9DGJ2	ID	Q9DGJ2 ; PRELIMINARY;	PRT;	147 AA.
AC	DR	Q9DGJ2 ;	RC	TISSUE=Skeletal muscle;
DT	DT	01-MAR-2001 (TREMBLrel. 16, Created)	RA	RX MEDLINE=21146139; PubMed=11247835;
DT	DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	RA	RA Marcinek D.J.; Bonaventura J.; Wittenberg J.B.; Block B.A.;
DT	DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	RT	RT "Oxygen affinity and amino acid sequence of myoglobins from endothermic and ectothermic fish.";
DE	DR	Myoglobin.	RL	RL Am. J. Physiol. Regul. Integr. Comp. Physiol. 280:R1123-R1133 (2001).
OC	DR	Thunnus alalunga (Albacore).	DR	DR EMBL: AF291837; AAG021111; - .
OC	DR	Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostei; Neoteleostei; Teleostei; Euteleostei; Neoteleostei; Buteleoste;	DR	HSSP: P02205; 1 MYT.
OC	DR	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Buteleoste;	DR	GO; GO:0005344; P: oxygen transporter activity; IEA.
OC	DR	InterPro; IPR002335; Myoglobin.	DR	GO; GO:0015671; P: oxygen transport; IEA.
OC	DR	Pfam: PF00042; Globin_1.	DR	InterPro; IPR002335; Myoglobin.
OC	DR	PRINTS: PR0613; MYOGLOBIN.	DR	PROSITE; PS01033; GLOBIN_1.
OC	DR	PROSITE; PS01033; GLOBIN_1.	DR	DR Pfam: PF00042; Globin_1.

Query Match	34.2%	Score 271.5; DB 13; Length 141;		
Best Local Similarity	42.3%	Pred. No. 1..6e-17;		
Matches	60; Conservative 20; Mismatches 59; Indels 3; Gaps 2;			
		DT 01-FEB-1997 (TREMBLrel. 02; Created)		
		DT 01-FEB-1997 (TREMBLrel. 02; Last sequence update)		
		DT 01-OCT-2003 (TREMBLrel. 25; Last annotation update)		
		DE Myoglobin (Fragment).		
		OS Notothenia coriiceps (black rockcod).		
		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Neoteleostei; Neoteleosteini;		
		OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleosteini;		
		OC Acanthomorpha; Acanthopterygii; Perciformes; Notothenioidei; Notochenidae; Notothernia.		
		OC Notothenioidei; Notochenidae; Notothernia.		
		OX NCBI_TaxID:8208;		
		RN [1]		
		RP SEQUENCE FROM N.A.		
		RC TISSUE-Heart ventricle;		
		RA Vayda M.E.; Small D.J.; Sidell B.D.; "Expression of the myoglobin gene in Antarctic channichthyid fishes." (In) Battaglia B., Valencia J., Walton S.W.H. (eds.); VI SCAR ANTARCTIC COMMUNITIES, pp.1-1, Cambridge University Press, Cambridge, UK (1996).		
		RN [2]		
		RP SEQUENCE FROM N.A.		
		PC TISSUE-Heart ventricle;		
		RX MEDLINE=98160008; PubMed=2694664;		
		RA Small D.J.; Vayda M.E.; Sidell B.D.; "A novel vertebrate myoglobin gene containing three A+T-rich introns is conserved among Antarctic teleost species which differ in myoglobin expression." (In) Battaglia B., Valencia J., Walton S.W.H. (eds.); VI SCAR ANTARCTIC COMMUNITIES, pp.1-1, Cambridge University Press, Cambridge, UK (1996).		
		RN [3]		
		RP SEQUENCE FROM N.A.		
		PC TISSUE-Heart ventricle;		
		RX MEDLINE=98160008; PubMed=2694664;		
		RA Small D.J.; Vayda M.E.; Sidell B.D.; "A novel vertebrate myoglobin gene containing three A+T-rich introns is conserved among Antarctic teleost species which differ in myoglobin expression." (In) Battaglia B., Valencia J., Walton S.W.H. (eds.); VI SCAR ANTARCTIC COMMUNITIES, pp.1-1, Cambridge University Press, Cambridge, UK (1996).		
		RN [4]		
RESULT 11				
Q9DGII9	PRELIMINARY;	PRT; 147 AA.		
AC Q9DGII9				
DB 01-MAR-2001 (TREMBLrel. 16; Created)				
DB 01-OCT-2003 (TREMBLrel. 25; Last sequence update)				
DE Myoglobin.				
OS Scomber japonicus (Chub mackerel).				
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Neoteleosteini; Scombroidei; Scombridae; Scomber.				
NCBI_TaxID=13676;				
RN [1]		RP SEQUENCE FROM N.A.		
		RC TISSUE-Skeletal muscle;		
		RX MEDLINE=21146139; PubMed=11247835;		
		RA Marcinek D.J.; Bonaventura J.J.; Wittenberg J.B.; Block B.A.; "Oxygen affinity and amino acid sequence of myoglobins from endothermic and ectothermic fish."		
		RT Am. J. Physiol. Regul. Integr. Comp. Physiol. 280:R1123-R1133 (2001).;		
		DR EMBL: AF291835; AAG02109_1; -.		
		DR HSSP; P02205; IMGT.		
		DR GO; GO:0005344; F:oxygen transporter activity; IEA.		
		DR GO; GO:015671; P:oxygen transport; IEA.		
		DR InterPro; IPR002335; Myoglobin.		
		DR Pfam; PF00042; Globin.		
		DR PRINTS; PR00613; MYOLOBIN.		
		DR PROSITE; PS01033; GLOBIN.		
		KW Heme; Oxygen transport; Transport.		
		SQ SEQUENCE 147 AA; 15766 MW; 44E9A4611EE0366 CRC64;		
		DR		
		RESULT 13		
Q9BNB0	PRELIMINARY;	PRT; 190 AA.		
AC Q9BNB0				
DB 01-MAR-2003 (TREMBLrel. 23; Created)				
DB 01-JUN-2003 (TREMBLrel. 23; Last sequence update)				
DE Mus musculus cyoglobin.				
GN CYGB				
OS Mus musculus (Mouse).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Neoteleosteini;				
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
NCBI_TaxID=10090;				
RN SEQUENCE FROM N.A.				
RC STRAIN=C57BL/6J; TISSUE-EYE;				
RX MEDLINE=22354683; PubMed=12466851;				
RA The RANTOM Consortium,				
RA the RIKEN Genome Exploration Research Group Phase I & II Team;				
RT "Analysis of the mouse transcriptome based on functional annotation of				
RESULT 12				
Q98963	PRELIMINARY;	PRT; 110 AA.		
ID Q98963				
AC Q98963;				

GenCore version 5.1.6 Copyright (c) 1993 - 2004 Compugen Ltd.	85	9.1	2599	3	AAV75098 Neisseria Abu38184 Protein e
OM protein - protein search, using sw model	85	9.1	2703	6	ABU38184 Protein e
Run on: August 10, 2004, 15:18:30 ; Search time 31.8547 Seconds	27	8.9	315	6	ABU32037 Protein e
Perfect score: 933	28	8.9	712	5	ABB91977 Herbicida
Sequence: 1 MSNDNTLVTDVRNGIDGH.....DELVARFLPMLKLTDFQI 184	29	8.9	1254	6	ABU25854 Protein e
Title: BLOSUM62	30	8.9	4384	6	ABU02251 S. pneumo
Scoring table: Gapext 10.0 , Gapext 0.5	31	8.9	6	ABU02251 Drosophil	
Searched: 1586107 seqs, 282547505 residues	32	8.9	319	4	ABU62080 Drosophil
Total number of hits satisfying chosen Parameters:	33	8.9	403	6	Abu34054 Protein e
Minimum DB seq length: 0	34	8.8	695	6	ABU41908 Protein e
Maximum DB seq length: 2000000000	35	8.8	831	4	AAV91714 C. glutam
Post processing: Maximum Match 0% Listing first 45 summaries	36	8.8	928	6	Abu33824 Protein e
Database : A_Geneseq_29Jan04:*	37	8.8	329	6	Abr56445 Bacteriop
1: geneseqp1980s:*	38	8.8	329	6	Abr44548 Bacteriop
2: geneseqp1190s:*	39	8.8	329	7	Add24123 Bacteriop
3: geneseqp2000s:*	40	8.8	330	5	ABG94315 RNA phage
4: geneseqp2001s:*	41	8.8	1333	6	Abu27622 Protein e
5: geneseqp2002s:*	42	8.8	2799	6	ABU37640 Protein e
6: geneseqp2003as:*	43	8.7	248	5	Abu41460 Human ORF
7: geneseqp2003bs:*	44	8.7	248	5	ABP04390 Human ORF
8: geneseqp2004s:*	45	8.7	1178	6	ABJ26182 Aspergill
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	ALIGNMENTS				
RESULT 1	XX				
ID AAE04641	XX				
AC AAE04641;	XX				
DT 11-SEP-2003 (revised)	XX				
DT 04-SEP-2001 (first entry)	XX				
XX Halobacterium salinarum HemAT-Hs protein.	XX				
KW Haem binding protein; HemAT-Hs; HemAT-Bs; gaseous ligand sensor; oxygen storage; artificial photosynthesis; signalling function; alpha-haemoglobin; myoglobin; therapy.	XX				
KW XX Halobacterium salinarum.	OS				
XX WO200140475-A2.	XX				
PD 07-JUN-2001.	XX				
PP 05-DEC-2000; 2000WO-US033048.	XX				
PR 06-DEC-1999; 99US-00455978.	XX				
PA (UYH-) UNIV HAWAII.	XX				
PI Alam M, Larsen R;	XX				
XX WPI; 2001-374832/39.	DR				
DR N-PSDB; AAD08991.	XX				
PS Claim 6; Page 10; 94pp; English.	XX				
SUMMARIES	XX				
Result No.	Query	Match	Length	DB ID Description	
- - -	- - -	- - -	- - -	- - -	
1	933	100.0	489	4 Aae0641 Halobacte	
2	203	21.8	39	4 Aae04677 Haem Prot	
3	142	15.2	432	4 Aae0642 Bacillus	
4	99	10.6	448	1 AAP70493 Protein G	
5	99	10.6	448	1 AAP95030 Protein G	
6	99	10.6	448	2 AAR07013 Protein G	
7	99	10.6	448	2 AAR10004 Streptoco	
8	99	10.6	448	2 MAR53290 Streptoco	
9	99	10.6	1175	2 AY34570 Porphory	
10	99	10.6	1232	2 AY34569 Porphory	
11	99	10.6	1266	2 AY34568 Porphory	
12	99	10.6	1269	2 AY34428 Porphory	
13	98	10.5	593	2 AAR07014 Protein G	
14	98	10.5	593	2 AAR62944 Streptoco	
15	98	10.5	594	2 AAR10005 Streptoco	
16	89	9.5	955	2 AAR57365 K39 polyp	
17	89	9.5	955	2 AAW03691 Leishmani	
18	88	9.4	505	6 ABU38101 Protein e	
19	87	9.3	496	6 Abu37440 Protein e	
20	87	9.3	505	6 ABP79309 N. gonorr	
21	87	9.3	523	2 AY34454 Porphory	
22	87	9.3	533	2 AY34330 Porphory	
23	86.5	9.3	318	3 AAB07677 Amino aci	
24	86	9.2	483	7 ADC31377 Human nov	
25	85.5	9.2	302	6 ADA33559 Acinetoba	
CC The present invention relates to isolated archael and bacterial haem binding protein, HemAT-Hs or HemAT-Bs which reversibly binds oxygen with low affinity. Haem binding protein is useful for controlled storage of oxygen by allowing haem binding protein to bind and store oxygen, and triggering the release of oxygen from haem binding protein by activating the signalling domain. Haem binding protein is useful for sensing gaseous ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem binding protein is useful for treating a patient suffering from low blood levels by administering and regulating the oxygen binding of the haem-binding protein by modifying the signalling domain. Haem binding protein	CC				

CC is useful for haem-based catalysis, for artificial photosynthesis and for identifying potential signalling functions of mutated alpha-haemoglobin CC and myoglobin causing several diseases. The present sequence is CC Halobacterium salinarium HemAT-Hs protein which is salt tolerant. CC (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 489 AA;

Query Match 100.0%; Score 933; DB 4; Length 489;
Best Local Similarity 100.0%; Pred. No. 2.6e-89;
Matches 184; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MSNDNTLTVTAVRNGIDGHALADRIGIDDEAIAWRLSFTGIDDDTMALAAEQPIFEAT 60
Db 1 MSNDNTLTVTAVRNGIDGHALADRIGIDDEAIAWRLSFTGIDDDTMALAAEQPIFEAT 60
Qy 61 ADALVTDYDFHLSYRTQDLFANSTKTEVQLKEHQYDTEYAAQRARIG 120
Db 61 ADALVTDYDFHLSYRTQDLFANSTKTEVQLKEHQYDTEYAAQRARIG 120
Qy 121 KHDVLGLGDPVYLGATRYTGLDALLADDVADREEAANADELVARFLPMLKLTF 180
Db 121 KHDVLGLGDPVYLGATRYTGLDALLADDVADREEAANADELVARFLPMLKLTF 180
Qy 181 DQO1 184
Db 181 DQO1 184

RESULT 2
AAE04677
ID AAE04677 standard; peptide; 39 AA.

XX AAE04677;

XX DT 11-SEP-2003 (revised)

DT 04-SEP-2001 (first entry)

XX DE Haem protein related Halobacterium salinarium protein fragment.
XX KW Haem binding protein; HemAT-Hs; HemAT-Bs; Gaseous ligand sensor;
KW oxygen storage; artificial photosynthesis; signalling function;
KW alpha-haemoglobin; myoglobin; therapy.
XX OS Halobacterium salinarium.
XX PR Location/Qualifiers
FH 19...39
FT /label= M2_Box
XX WO200140475-A2.

XX Key Domain
PR 06-DEC-1999; 99US-00455978.
XX PA (UYHA-) UNIV HAWAII.
XX PI Alam M, Larsen R;

XX PS XX

XX WPI; 2001-374832/39.

XX DR 07-JUN-2001.

XX PP 05-DEC-2000; 2000WO-US033048.

XX PR 06-DEC-1999; 99US-00455978.

XX PA (UYHA-) UNIV HAWAII.

XX Alam M, Larsen R;

XX WPI; 2001-374832/39.

XX Novel isolated bacterial heme binding protein, HemAT-Hs or HemAT-Bs which

PT reversibly binds oxygen with low affinity, useful for controlled storage

PR of oxygen and for sensing gaseous ligands such as oxygen.

XX Disclosure; Page 16; 94pp; English.

XX The present invention relates to isolated archaeal and bacterial haem

CC binding protein, HemAT-Hs or HemAT-Bs which reversibly binds oxygen with

CC low affinity. Haem binding protein is useful for controlled storage

CC of oxygen by allowing haem binding protein to bind and store oxygen, and

CC triggering the release of oxygen from haem binding protein by activating

CC the signalling domain. Haem binding protein is useful for sensing gaseous

CC ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem

CC binding protein is useful for treating a patient suffering from low blood

CC levels by administering and regulating the oxygen binding domain. Haem binding protein

CC is useful for haem-based catalysis, for artificial photosynthesis and for

CC identifying potential signalling functions of mutated alpha-haemoglobin

CC oxygen by allowing haem binding protein to bind and store oxygen, and
CC triggering the release of oxygen from haem binding protein by activating
CC the signalling domain. Haem binding protein is useful for sensing gaseous
CC ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem
CC binding protein is useful for treating a patient suffering from low blood
CC levels by administering and regulating the oxygen binding domain. Haem binding protein
CC is useful for haem-based catalysis, for artificial photosynthesis and for
CC identifying potential signalling functions of mutated alpha-haemoglobin
CC protein related Halobacterium salinarium protein fragment. (Updated on 11-
CC SEP-2003 to standardise OS field)

XX Sequence 39 AA;

SQ Query Match 21.8%; Score 203; DB 4; Length 39;
Best Local Similarity 100.0%; Pred. No. 6e-14;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 96 QAEYLLGLGRGYDTEYAAQRARIGKTHDVUGLGPDVYL 134
Db 1 QAEYLLGLGRGYDTEYAAQRARIGKTHDVUGLGPDVYL 39

RESULT 3

AAE04642 standard; protein; 432 AA.

ID AAE04642;

XX AC AAE04642;

XX DT 04-SEP-2001 (first entry)

XX Bacillus subtilis HemAT-Bs protein.

XX DE Bacillus subtilis.

XX KW Haem binding protein; HemAT-Hs; HemAT-Bs; Gaseous ligand sensor;

KW oxygen storage; artificial photosynthesis; signalling function;

KW alpha-haemoglobin; myoglobin; therapy.

XX OS Bacillus subtilis.

XX XX WO200140475-A2.

XX PN 07-JUN-2001.

PD 05-DEC-2000; 2000WO-US033048.

XX XX 99US-00455978.

XX PA (UYHA-) UNIV HAWAII.

PA Alam M, Larsen R;

PR XX WPI; 2001-374832/39.

DR N-FSDB; AAD08992.

XX PT Novel isolated bacterial heme binding protein, HemAT-Hs or HemAT-Bs which

PT reversibly binds oxygen with low affinity, useful for controlled storage

PR of oxygen and for sensing gaseous ligands such as oxygen.

XX Disclosure; Page 11; 94pp; English.

XX The present invention relates to isolated archaeal and bacterial haem

CC binding protein, HemAT-Hs or HemAT-Bs which reversibly binds oxygen with

CC low affinity. Haem binding protein is useful for controlled storage

CC of oxygen by allowing haem binding protein to bind and store oxygen, and

CC triggering the release of oxygen from haem binding protein by activating

CC the signalling domain. Haem binding protein is useful for sensing gaseous

CC ligands such as oxygen, NO, CO, or CN. Blood substitute comprising haem

CC binding protein is useful for treating a patient suffering from low blood

CC levels by administering and regulating the oxygen binding domain. Haem binding protein

CC is useful for haem-based catalysis, for artificial photosynthesis and for

CC identifying potential signalling functions of mutated alpha-haemoglobin

CC and myoglobin causing several diseases. The present sequence is Bacillus subtilis HemAT-Bs protein

Sequence 432 AA;	Query Match 15.2%; Score 142; DB 4; Length 432;	CC non-pathogenic host. Suitable cloning vectors are lambda gt11, M13mp9 and PGX1066. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to correct OS field.)
Best Local Matches 39; Conservative 29; Mismatches 76; Indels 10; Gaps 2;	Best Local Similarity 25.3%; Pred. No. 4.8e-06;	CC
Qy 31 ABIAWRLSFTGIDDMMALAAEQLPFEATAALVTDYDLYESYERTQDLPANSKTKTE 90	Qy 10 TADVRGIDGHIALDRIGLDEAEIAANRSLSGFTGIDD----DTMALARAE-----	CC
Db 32 ADVKQKLMVRQGDAELYVLEQPLQIENIVNIDFAFYKNDVILSMIDLI-NDHSVVD 90	Db 41 TPILRNGC---LTNLIGNSSETTLARNEESATADLTAAAYADTVAAAGAAAWEA 97	CC
Qy 91 QLKETQAEYLGLGRGEYDTEAQQRARIKGKLYHDVLGLGPDVYLGAVTRYTGLLDALAD 150	Qy 54 ---QPLFEATADAL-----VTDFYDHLESYERTQDLPANSKTKTEOLKETQAEYLGL 102	CC
Db 91 RLKQTLKRHICQEMFAGVYDDEFIEKRNRIASIHLRIGLIPKXNMGAFOELLMSMIDIY- 148	Db 98 AAAADALAKAARDALKEPNKYGVSDYKNL-----INNAKTYEGIKDLQAQNV-- 145	CC
Qy 151 DVVADRGEEAAAAYADELVAFARFLPMLKLTFDQI 184	Qy 103 LGRGEYDTEYIAYAQRATIGKTHDVLGLGPDVYLGAYTRYTGSHLDALADDVYADRGEAAA 162	CC
Db 149 -----EASITNQELLIKAIAKATTKILNLBEQQL 175	Db 146 -----ESARKARISEATD-----GLSDFLKSQTBA--EDTVK 175	CC
RESULT 4	Qy 163 AVDELVYARFL 172	CC
AAP70493	Db 176 SIELABRAKVL 185	CC
ID AAP70493 standard; protein; 448 AA.		
XX		
AC AAP70493;		
XX DT 27-AUG-2003 (revised)	RESULT 5	CC
XX DT 25-MAR-2003 (revised)	AAP95030	CC
XX DT 06-MAR-1991 (first entry)	ID AAP95030 standard; protein; 448 AA.	CC
DB Protein G.	XX AC AAP95030;	CC
XX KW antibody; Fc receptor.	XX AC AAP95030;	CC
OS Streptococcus sp; 'Lancefield Group G strain'.	XX DT 25-MAR-2003 (revised)	CC
XX OS	DT 04-JUL-1990 (first entry)	CC
XX Key Location/Qualifiers	XX DB Protein G.	CC
FH Active-site FT 228..297	XX KW Protein G; immunoglobulin; Fc receptor; ds.	CC
FT /label= active site B1	XX OS Streptococcus sp.	CC
FT 298..352	XX PN WO8B10306-A.	CC
FT /label= active site B2	XX PD 29-DEC-1998.	CC
XX PN WO8705025-A.	XX PF 20-JUN-1988;	CC
XX PD 27-AUG-1987.	XX PR 19-JUN-1987;	CC
XX PF 17-FEB-1987; 87WO-US000329.	XX PA (GENEX) GENEX CORP.	CC
XX PR 14-FEB-1996; 86US-00822354.	PA (PHAA) PHARMACIA AB.	CC
PR 23-APR-1986; 86US-00854887.	XX PI Fahnestock SR;	CC
XX (GENEX) GENEX CORP.	XX DR WPI:1989-023848/03.	CC
PA (FAHN/) FAHNESTOCK S R.	XX N-PSDB; AAN9193.	CC
XX PI Fahnestock SR;	XX N-PSDB; AAN9193.	CC
XX DR 1987-250197/35.	XX Cloned protein G variant genes - expressing proteins having PT immunoglobulin-binding properties of protein G and derived from PT streptococcus sp.	CC
XX DR N-PSDB; AAN70811.	XX Disclosure; Page; 116pp; English.	CC
XX Cloned Protein G gene - used for producing Protein G for detection and PT purification of antibodies and treatment of diseases.	XX Protein G of non-pathogenic streptococcus and variants may be isolated, CC useful as bacterial Fc receptors eg in purification and detection of Abs.	CC
PT PT screening of hybridoma clones and treatment of disease. (Updated on 25- MAR-2003 to correct PA field.)	CC	CC
PT XX Sequence 448 AA;	XX Sequence 448 AA;	CC
PS Query Match	10.6%; Score 99; DB 1; Length 448;	CC

Best Local Similarity	25.3%	Pred.	No. 0.17;	Score 99;	DB 2;	Length 448;
Matches	48;	Conservative	25;	Mismatches	45;	Indels
Qy	10 TADVRNGIDGHALADRIGLDEAEIAWRSLSTFGIDD-----DTWAALAAE-----53	Gaps	8;	Best Local Similarity	25.3%;	Pred.
Db	41 TPILRNGE--LTNLIGNSSETTLAARNNEESATADLTAAAVADTVAAENAGAAWEA 97	Matches	48;	Conservative	25;	Mismatches
Qy	54 ---QPLFEATADL-----VTDPEYDHLSSEYERTQDLFANSTKTVFQKETOAEYLG 102	Qy	10 TADVRNGIDGHALADRIGLDEAEIAWRSLSTFGIDD-----DTWAALAAE-----53	Best Local Similarity	25.3%;	Pred.
Db	98 AAAADALAKAADALEKFNKYGVSDYKNL-----INNAKTVEGIKDLOAQNV--145	Db	41 TPILRNGE--LTNLIGNSSETTLAARNNEESATADLTAAAVADTVAAENAGAAWEA 97	Matches	48;	Conservative
Qy	103 LGRGEYDTEYAAQRARIGKHDVLGLGPDVYLGATRYTGGDLADDVVAADRGEAAA 162	Qy	54 ---QPLFEATADL-----VTDPEYDHLSSEYERTQDLFANSTKTVFQKETOAEYLG 102	Best Local Similarity	25.3%;	Pred.
Db	146 -----ESAKKARISEATD-----GLSDFELKSQTPA---EDTVK 175	Db	98 AAAADALAKAADALEKFNKYGVSDYKNL-----INNAKTVEGIKDLOAQNV--145	Matches	48;	Conservative
Qy	163 AVDELVARFL 172	Qy	103 LGRGEYDTEYAAQRARIGKHDVLGLGPDVYLGATRYTGGDLADDVVAADRGEAAA 162	Best Local Similarity	25.3%;	Pred.
Db	176 SIELAAEKVL 185	Db	146 -----ESAKKARISEATD-----GLSDFELKSQTPA---EDTVK 175	Matches	48;	Conservative
RESULT 6						
AAR07013	standard; protein; 448 AA.	AAR10004	standard; protein; 448 AA.			
ID	AAR07013	ID	AAR10004			
XX		XX				
AC	AAR07013;	AC	AAR10004;			
XX		XX				
DT	24-OCT-2003 (revised)	DT	24-OCT-2003 (revised)			
DT	25-MAR-2003 (revised)	DT	25-MAR-2003 (revised)			
DT	17-JAN-1991 (first entry)	DT	13-MAR-1991 (first entry)			
DB	Protein G variant with two active sites.	DB	Protein G variant with two active sites.			
XX		XX				
KW	Immunglobulin.	KW	Immunglobulin.			
XX	Streptococcus sp; Lancefield Group G strain.	XX	Streptococcus sp; GX7809.			
OS		OS				
XX		XX				
RESULT 7						
AAR07013	Location/Qualifiers	AAR10004	Location/Qualifiers			
XX		XX				
AC		AC				
XX		XX				
DT		DT				
DT		DT				
DT		DT				
XX		XX				
DB		DB				
XX		XX				
KW		KW				
XX		XX				
FT	Key	FT	Key			
FT	Active-site	FT	Active-site			
FT	228; /label= B1	FT	228; /label= Active site B1			
FT	298; .352	FT	298; .352			
FT	/label= B2	FT	/label= Active site B2			
XX		XX				
PN	US4956296-A.	PN	US4977247-A.			
XX		XX				
PD	11-SEP-1990.	PD	11-DEC-1990.			
XX		XX				
PP	20-JUN-1988;	PP	88US-00209236.			
XX		XX				
PR	14-FEB-1986;	PR	19-MAY-1989;	89US-00354264.		
PR	88US-00829354.	PR	14-FEB-1986;	86US-00829354.		
PR	00854887.	PR	23-APR-1986;	86US-00854887.		
PR	87WO-US000329.	PR	17-FEB-1987;	87WO-US000329.		
PR	19-JUN-1987;	PR	19-JUN-1987;	87US-0063959.		
XX	88US-00063959.	PR	20-JUN-1988;	88US-00209236.		
PA	(GEMX) GENEX CORP.	PA	(GEMX) GENEX CORP.			
XX	Fahnestock SR;	XX	Fahnestock SR, Lee T, Wroble MH;			
DR	WPI; 1990-297491/39.	DR	WPI; 1991-006758/01.			
XX	N-PSDB; AAQ06018.	XX	WPI; 1991-006758/01.			
PT	Recombinant Protein G variants - obtd. using a cloned gene encoding	PT	N-PSDB; AAQ10001.			
PT	Protein G from Streptococcus sp., used for binding immunoglobulin.	PT	Protein G variants - used for detection, isolation and			
XX		XX	purification, immunoglobulin(s) and immunoglobulin fragments.			
PS	Disclosure; Fig 8a-c; 48pp; English.	PS	Disclosure; Fig 8; 52pp; English.			
XX		XX				
CC	Fragments and variants of the sequence are claimed esp. where	CC	Protein G gene product may be modified allowing the variant to be			
CC	incorporated into a non-pathogenic host eg. E.coli, and expressed at high	CC	immobilised and exhibit different binding profiles. The bound protein is			
CC	levels. The variants have a higher binding efficiency and capacity for	CC	immobilised and purification and detection of IgG and fragments. (Updated on 24-			
CC	immunoglobulin, and may be used for purifying, detecting and isolating	CC	OCT-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI			
CC	antibodies. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS field)	CC	field.) (Updated on 24-OCT-2003 to standardise OS field)			
SQ	Sequence 448 AA;	SQ				

XX Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
PT Hocking DM, Webb EA;
XX WPI; 1999-385613/32.
PS N-PSDD; AAX9178.
XX Antigenic Porphyromonas gingivalis peptides for preventing gingivitis.
XX Claim 1; Page 563-565; 58app; English.
CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY3483. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can be
CC used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially Gingivitis. (Updated on 20-MAR-2003 to correct PR field.)
XX (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 1175 AA;
SQ Query Match 10.6%; Score 99; DB 2; Length 1175;
Best Local Similarity 25.6%; Pred. No. 0.66;
Matches 23; Mismatches 47; Indels 64; Gaps 9;
SQ Sequence 1232 AA;
Query Match 10.6%; Score 99; DB 2; Length 1232;
Best Local Similarity 25.6%; Pred. No. 0.71;
Matches 46; Conservative 23; Mismatches 47; Indels 64; Gaps 9;
Qy 16 GIDGHALADRI-----GLDEABIAWRLSFTG----IDDITMAALLAEEQPHI--- 56
Db 333 GLDGLASITRLSLRRNQISKLGDLRKLVDGSNDIQSIDDIKLIALPLEQTEKL 392
Qy 57 -----PEATADALVTDYDHLSEYERTQDLEANSKTKVEQLKETQAEY-----LIG- 102
Db 393 RTHDNPPVASSGLLSPDNH-----PEIKALLEKEKEKQKRTTSVBYHPFCKVMILLGN 446
Qy 103 --LGR--GENDTEYAQRARIGKIHIVGL-----GPDVTLGAYTRYT 142
Db 447 HSSGKTTPLSQIDNTNYQ-----KNTVLSIHSRSNNPNAIFYDFGGQDYHGIYQAFPT 501
Qy 103 --LGR--GENDTEYAQRARIGKIHIVGL-----GPDVYLGAATRYT 142
Db 504 HSSGKTTPLSQIDNTNYQ-----KNTVLSIHSRSNNPNAIFYDFGGQDYHGIYQAFPT 558
RESULT 11
AAV34568 standard; protein: 1266 AA.
XX AAY34568;
AC AAY34568;
DT 27-AUG-2003 (revised)
DT 20-MAR-2003 (revised)
DT 25-AUG-1999 (first entry)
XX DE (revised)
XX DT 27-AUG-2003 (revised)
XX DT 20-MAR-2003 (revised)
XX DT 25-AUG-1999 (first entry)
DB Porphyromonas gingivalis protein PG9.
KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis; vaccine;
KW antigenic.
OS Porphyromonas gingivalis.
XX WO929870-A1.
XX 17-JUN-1999.
PD 10-DEC-1998; 9BWO-AU0001023.
XX 10-DEC-1998; 9BWO-AU0001023.
PR 10-DEC-1997; 9BAU-00000839.
PR 31-DEC-1997; 9BAU-00001182.
PR 30-JAN-1998; 9BAU-00001546.
PR 10-MAR-1998; 9BAU-00002264.
PR 09-APR-1998; 9BAU-00002911.
PR 23-APR-1998; 9BAU-00003128.
PR 05-MAY-1998; 9BAU-00003338.
PR 22-MAY-1998; 9BAU-00003654.
PR 09-JUL-1998; 9BAU-00004917.
PR 30-JUL-1998; 9BAU-00004963.
PR 04-AUG-1998; 9BAU-00005028.
XX (CSLC-) CSL LTD.
XX PR 22-JUL-1998; 9BAU-00004917.
XX PR 30-JUL-1998; 9BAU-00004963.
XX PR 04-AUG-1998; 9BAU-00005028.

PD 11-SEP-1990.
 XX 20-JUN-1988; 88US-00209336.
 PP 86US-00829354.
 PR 14-FEB-1986; 86US-00854387.
 PR 23-APR-1986; 87WO-US000329.
 PR 17-FEB-1987; 87US-00063959.
 PR 19-JUN-1987; 87US-00063959.
 XX (GEMX) GENEX CORP.
 PA DR; AAQ06019.
 XX Fahnstock SR;
 XX Disclosure: Fig 9; 48pp; English.
 XX PT Recombinant Protein G variants - obt'd. using a cloned gene encoding
 CC Protein G from Streptococcus sp., used for binding immunoglobulin.
 XX PS Disclosure: Fig 9; 48pp; English.
 XX Fragments and variants of the sequence are claimed esp. where
 CC incorporated into a non-pathogenic host eg. E.coli, and expressed at high
 CC levels. The variants have a higher binding efficiency and capacity for
 CC immunoglobulin, and may be used for purifying, detecting and isolating
 CC antibodies. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-
 CC OCT-2003 to standardise OS field)
 XX Sequence 593 AA;
 SQ Query Match 10.5%; Score 98; DB 2; Length 593;
 Best Local Similarity 25.3%; Pred. No. 0.32; Gaps 8;
 Matches 48; Conservative 25; Mismatches 45; Indels 72;
 Qy 10 TADYRNGIDGHALADRIGLDEAIAWRLSFTGIDD----DTMAAIAAE----- 53
 Db 41 TPTRNGBEE--LTNLIGNSSETTLARNNEESATADLTAAVADTVAAARENAGAAWEA 97
 Qy 54 ---QPLISREATADL---VTDPEVDHLESEVRTQLIFANSTKTVQLKETQAEYLIG 102
 Db 98 AAAADALAKAKADALKKEFPNKYGSDYXKLN-----INNAKTYEGVKDLOAQTV-- 145
 Qy 103 LGREGYDTEYAAQRARTGKHDVGLGDVLYGAYTRVYGLDADVVADRGEEAA 162
 Db 146 -----ESAKKAR1SEATD-----GSDFEIKSQTPA--EPTVK 175
 XX PR 21-APR-1992; 92US-00871539.
 Qy 163 AYDELVARFL 172
 Db 176 SIBLAEEAKVL 185
 RESULT 14
 ID AAR62944 standard; protein; 593 AA.
 XX AC AAR62944;
 AC XX DT 25-MAR-2003 (revised)
 DT 10-JAN-1995 (first entry)
 XX DB Streptococcus Protein G derived from strain GX7805.
 XX KW Streptococcus Protein G; variant; IgG binding activity; immunoglobulin;
 KW Lancefield Group G; bacterial Fc receptor.
 XX OS Streptococcus sp. GX7805.
 XX FH Location/Qualifiers
 Domain 106..140
 /label= A1
 Region 141..178
 /label= a1
 FT /note= "linking region"
 FT
 FT Domain
 /label= A2
 FT Region
 /label= a2
 /note= "linking region"
 FT Domain
 /label= A3
 FT Misc-difference 269
 FT /note= "Corresponds to GGA codon"
 FT Active-site 303..357
 FT /label= B1
 /note= "involved in IgG binding activity"
 PI Region 358..372
 /label= b
 /note= "linking region"
 PA Active-site 373..427
 FT /label= B3
 /note= "involved in IgG binding activity"
 PT Region 428..442
 /label= b
 /note= "linking region"
 FT Active-site 443..497
 FT /label= B2
 /note= "involved in IgG binding activity"
 FT Region 466..531
 /note= "corresponds to ACT codon"
 FT Region 531..535
 /label= C1
 FT Region 536..540
 /label= C2
 FT Region 541..545
 /label= C3
 FT Region 546..550
 /label= C4
 FT Region 551..555
 /label= C5
 FT Misc-difference 592
 /note= "corresponds to GAA codon"
 PR US5312901-A.
 XX PD 17-MAY-1994.
 XX PR 19-JUN-1990; 90US-00540169.
 XX PA (PHAA) PHARMACIA LKB BIOTECHNOLOGY AB.
 XX PA Fahnestock SR;
 XX PI Fahnestock SR;
 XX DR WPI; 1994-159179/19.
 DR N-PSDB; AAQ75036.
 XX PT New recombinant streptococcal protein G variants - useful for antibody
 detection and purification and for therapy.
 XX PS Example 5; Fig 9; 48pp; English.
 XX CC A 2.4kb HindIII fragment containing the entire coding sequence for
 CC Protein G was isolated from Streptococcus GX7805 using the 1.9kb Protein
 CC G coding sequence from Streptococcus GX7809. The Protein G has IgG-
 CC binding activity which has been localised to the B repeating structure.
 CC Streptococcal Protein G variants comprising the B domains are claimed.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 593 AA;

Query Match. Similarity 10.5%; Score 98; DB 2; Length 593;
 Best Local Similarity 25.3%; Pred. No. 0.32; Mismatches 45; Indels 72; Gaps 8;
 Matches 48; Conservative 25; MisMatches 45; Indels 72; Gaps 8;

CC -MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI
 CC field.) (Updated on 24-OCT-2003 to standardise OS field)

Sequence 594 AA;

Query Match 10.5%; Score 98; DB 2; Length 594;
 Best Local Similarity 25.3%; Pred. No. 0.32; Mismatches 45; Indels 72; Gaps 8;
 Matches 48; Conservative 25; MisMatches 45; Indels 72; Gaps 8;

CC -MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI
 CC field.) (Updated on 24-OCT-2003 to standardise OS field)

Sequence 594 AA;

Qy 10 TADVRNGIDGHALADRIGIDEAERIAWRLSFTGIDD-----DTMAAIAAE---- 53
 Db 41 TP1RNGSE--LTNLIGSETTLALRNEESATADLTAAAVADTVAAAENAGAAWEA 97

Qy 54 ----QPLFEATADAL-----VTDIFDHLIESYERTQDLFANSTKTVTQLKETQAEMLLG 102
 Db 98 AAAADALAKAADKEFKNYGSDDYZKNL-----INNAKTYVGKVDLQAQVY- 145

Qy 103 LGRGEYDTEYAAQBARIGKTHDVGLGDPDVYLGAYTRYTGILDAADDVVADRGEEAA 162
 Db 146 -----ESAKKARISEATD-----GLSDFLKSQTPA--EDTVK 175

Qy 163 AVDELVARFL 172
 Db 176 SIELAAEKVL 185

RESULT 15
 AAR10005
 ID AAR10005 standard; protein; 594 AA.

XX AC AAR10005;
 XX DT 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 13-MAR-1991 (first entry)

XX DE Streptococcus GX7805 protein G.
 XX KW Immunoglobulins; Ig.
 XX OS Streptococcus sp; GX7805.

XX Key Location/Qualifiers
 FH Active-site 304..358
 FT /label= Active Site B1
 FT 374..428
 FT /label= Active Site B3
 FT 444..498
 FT /label= Active Site B2

XX PN US4977247-A.
 XX PD 11-DEC-1990.
 XX PP 19-MAY-1999; 89US-00354264.
 PR 14-FEB-1986; 86US-00829354.
 PR 23-APR-1986; 86US-00854887.
 PR 17-FEB-1987; 87WO-US000329.
 PR 19-JUN-1987; 87US-00063959.
 PR 20-JUN-1988; 88US-0020236.
 PA (GEMX) GENEX CORP.
 XX Fahnestock SR, Lee T, Wroble MH;
 XX DR WPI; 1991 006758/01.
 XX DR N-PSDB; AQ10002.

PT Immobilised protein G variants - used for detection, isolation and
 PT purification, immunoglobulin(s) and immunoglobulin fragments.
 XX Disclosure; Fig 9; 52pp; English.

CC Protein G gene product may be modified allowing the variant to be
 CC immobilised and exhibit different binding profiles. The bound protein is
 CC useful in purification and detection of Ig's and fragments. (Updated on 25

B
I
D
N
T

Result No.	Score	Query Match	Length	DB ID	Description	SUMMARIES	
1	91.5	9.8	955	2	US-08-428-414A-3	Sequence 3, Appli	
2	89	9.5	955	1	US-08-016-676B-1	Sequence 1, Appli	
3	89	9.5	955	1	US-08-282-845-2	Sequence 2, Appli	
4	89	9.5	955	5	PCT-US94-00324-1	Sequence 1, Appli	
5	85.5	9.2	302	4	US-09-328-552-4846	Sequence 4846, Appli	
6	83.5	8.9	319	4	US-09-489-039A-3872	Sequence 8872, Appli	
7	82	8.8	542	4	US-09-252-591A-21396	Sequence 21396, A	
8	81.5	8.7	1253	4	US-09-252-591A-30019	Sequence 30019, A	
9	80.5	8.6	700	4	US-09-252-591A-19384	Sequence 19384, A	
10	80	8.6	553	4	US-09-252-591A-32970	Sequence 32970, A	
11	80	8.6	677	4	US-09-252-591A-18102	Sequence 18102, A	
12	79	8.5	438	4	US-09-540-336-2295	Sequence 2595, Appli	
13	79	8.5	878	3	US-08-941-936-2	Sequence 2, Appli	
14	78.5	8.4	733	4	US-09-328-5539	Sequence 5539, Appli	
15	78.5	8.4	755	5	PCT-US93-01923-3	Sequence 3, Appli	
16	78.5	8.4	759	5	PCT-US93-01923-2	Sequence 2, Appli	
17	78.5	8.4	766	1	US-08-230-491A-3	Sequence 3, Appli	
18	78.5	8.4	766	1	US-08-619-380A-3	Sequence 3, Appli	
19	78.5	8.4	766	2	US-08-940-391-3	Sequence 3, Appli	
20	78.5	8.4	766	4	US-09-194-336-1	Sequence 1, Appli	
21	78.5	8.4	766	4	US-10-002-93-6	Sequence 6, Appli	
22	78	8.4	443	4	US-09-328-352-6943	Sequence 6943, Appli	
23	78	8.4	510	4	US-09-489-039A-11123	Sequence 11123, Appli	
24	77	8.3	297	4	US-09-252-591A-17492	Sequence 17492, A	
25	77	8.3	320	4	US-09-252-591A-21006	Sequence 21006, A	
26	76.5	8.2	579	4	US-09-252-591A-18063	Sequence 18063, A	
27	76	8.1	531	4	US-08-976-063E-34	Sequence 34, Appli	

Copyright (c) 1993 - 2004 Compugen Ltd.
GenCore version 5.1.6
OM protein - protein search, using sw model
Run on: August 10, 2004, 15:23:41 ; Search time 9.35593 Seconds
Perfect score: 933
Sequence: 1 MSNDNTLVTDVRNGIDGH.....DEVARFLPMLKLTDFDQQI 184
Title: BLOSM62
Scoring table: Gapext 0.0 , Gapext 0.5
Searched: 389414 seqs, 51625971 residues
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/2/iaa/5A_COMB_pep:
2: /cgn2_6/prodata/2/iaa/5B_COMB_pep:
3: /cgn2_6/prodata/2/iaa/6A_COMB_pep:
4: /cgn2_6/prodata/2/iaa/6B_COMB_pep:
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB_pep:
6: /cgn2_6/prodata/2/iaa/backfile1.pep:
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

ALIGNMENTS

ALIGNS

RESULT 1

US-08-428-414A-3

; Sequence 3, Application US/08428414A
; Patent No. 5912166
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF LEISHMANIASIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADRESSEEE: SBED and BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
CURRENT APPLICATION DATA:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
APPLICATION NUMBER: US/08/428,414A
FILING DATE: 21-APR-1995
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Kadlecik, Ann T.
REGISTRATION NUMBER: 39,244
REFERENCE/DOCKET NUMBER: 210121-407
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6131
TELEX: 3723836 SEEDANDBERRY
INFORMATION FOR SEQ ID NO.: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 955 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-428-414A-3

Query Match Best Local Similarity Score 9.8%; Length 955;
Matches 54; Conservative 22; Mismatches 77; Indels 45; Gaps 8;
QY 13 VRNGIDGHADRLGIDDEABIAWRLSFGIDDITMAAEQPFLFATADALVTT--DEFY 69
530 VRRRLDAELASEREKL-----ESTVAQLEERQREVALDALQTHQRKLQ 574
Db
QY 70 DHLEYERTQDLFANSKTVEQLKETQAELYLLGRGYDTE-YAAQRARIGKIHVDVLGL 128

RESULT 2
 US-08-006-676B-1 Application US/08006676B
 Patent No. 5411865

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
 TITLE OF INVENTION: A 210KD Antigen Present in Leishmania

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Apple Macintosh Operating System 7.1
 SOFTWARE: Microsoft Word for Macintosh 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/282,845

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/006,676

FILING DATE: JANUARY 15, 1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perkins, Patricia Anne

REGISTRATION NUMBER: 34,693

REFERENCE/DOCKET NUMBER: 5004-A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 955 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-282-B45-2

Query Match 9.5%; Score 89; DB 1; Length 955;
 Best Local Similarity 29.3%; Pred. No. 0.3; Mismatches 19; Indels 30; Gaps 7;
 Matches 49; Conservative 19; MisMatches 69; DelIns 30; Gaps 7;

Qy 44 DDTMAALAAEQQPLEATADALV-T--DFYDHLSEYERTDLFANSTKTVQQLKETQAYL 100
 Db 546 ESTAQLEREQREREVALIDQTHQKLOEALESSRTA--AERDQLIQLTQLOSE-R 601

Qy 101 LGIGRGEYDTE-YAAQRARIGKIHDLVGLGPDVYLGA---YTRYTYGHL----- 145
 Db 602 TQLSQVVTDRRLTRDLQRIQYEGETLARDVALCAAQMEARYHAAVPHLQTLLELAT 661

Qy 146 ---DALADDVVAADRGEAAAAYDELV---ARFLPMLKLITFQOQI 184
 Db 662 EWEEDALRERALAERDEAAAELDAASTSQNARESACERLTSLEQQL 708

RESULT 4
 PCT-US94-00324-1
 Sequence 1, Application PC/TUSS9400324
 GENERAL INFORMATION:
 APPLICANT: Reed, Steven
 TITLE OF INVENTION: Diagnosis of Leishmaniasis
 NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh

RESULT 3
 US-08-282-845-2
 Sequence 2, Application US/08282845
 Patent No. 5719263
 GENERAL INFORMATION:

OPERATING SYSTEM: Apple System 7.1
 SOFTWARE: Microsoft Word, version 5.1a
 CURRENT APPLICATION DATA:
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION NUMBER: PCT/US94/00324
 FILING DATE: 15-JAN-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Perkins, Patricia Anne
 REGISTRATION NUMBER: 34 693
 RECOMMENDATION/DOCKET NUMBER: 5004-WO
 TELEPHONE: (206) 567-0430
 TELEX/FAX: (206) 233-0644
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 955 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

PCT/US94-00324-1

RESULT 6
 US-09-489-039A-8872
 ; Sequence 8872, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE: 2709/2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO: 8872
 ; LENGTH: 319
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 ; US-09-489-039A-8872

Query Match 9.5%; Score 89; DB 5; Length 955;
 Best Local Similarity 29.3%; Pred. No. 0.3; Mismatches 19; Indels 30; Gaps 7;
 Matches 49; Conservative 51; Mismatches 69; Indels 30; Gaps 9;

Qy 44 DDTMALARAAEQPLFEATADALVY---DFYDHLLESYERTQDLPANSKTVKEQILKETQAELY 100
 Db 546 ESTVAQLEREQREVALDQTHQLRQLQEALESSERTA--AERDQLLQQITELOSE-R 601

Qy 101 LGLGRGEYDE-YAAQPARIGKIHDVYGLGPDVYLG---YTRYTYGGL---
 Db 602 TQLSQVYTDRERFLTRDQI QYEYGETELARVALCRAQEMARYHAFFHQLTLLBLAT 661

Qy 146 ---DALADDVYADRGEEAAAYDELV---ARFLPMLKLITTEDQI 184
 Db 662 EWEADAERALAERDEAAEEDAAASTSQNARESACERLTSLEQQL 708

Qy 149 ADDVVADRGEEAAAYDELVA 169
 Db 284 VDPGYAESEGRRAARQJIEQIA 304

RESULT 5
 US-09-328-352-4846

Sequence 4846, Application US/09328352

GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et. al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BAUMANNI FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328,352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO: 4846
 ; LENGTH: 302
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 ; US-09-328-352-4846

Query Match 9.2%; Score 85.5; DB 4; Length 302;

Best Local Similarity 29.4%; Pred. No. 0.14; Mismatches 23; Indels 41; Gaps 7;

Qy 25 RIGL---DEAIIJAWRLSFTGDDDTMAALAAEQPLFEATADALVTDYDLESTERQDL 81
 Db 179 RIGLNSVNDSAELIKSLP-QGPBNNTLVY---DPPYAKQDLYTFNHFQDHVEIMRAL 234

Query Match 9.2%; Score 85.5; DB 4; Length 302;

Best Local Similarity 29.4%; Pred. No. 0.14; Mismatches 23; Indels 41; Gaps 7;

Qy 82 FANSTK-----TVBQLKETQAELYLGLGRGEYDETEAAQRARIGK---IHIVLG 129
 Db 235 KSSSRNWIWYDNVDAIRETYKDFRV---LEYSIQLYTAQQKXIGEWIFNSNDV--LI 288

Query Match 8.8%; Score 82; DB 4; Length 542;

Best Local Similarity 25.3%; Pred. No. 0.84; Mismatches 49; Conservative 19; Indels 60; Gaps 8;
 ; SEQ ID NO: 21396
 ; LENGTH: 542
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-21396

Query Match 8.8%; Score 82; DB 4; Length 542;

Best Local Similarity 25.3%; Pred. No. 0.84; Mismatches 49; Conservative 19; Indels 60; Gaps 8;
 ; SEQ ID NO: 21396
 ; LENGTH: 542
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-21396

Query Match 8.8%; Score 82; DB 4; Length 542;

Best Local Similarity 25.3%; Pred. No. 0.84; Mismatches 49; Conservative 19; Indels 60; Gaps 8;
 ; SEQ ID NO: 21396
 ; LENGTH: 542
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-21396

Query Match 8.8%; Score 82; DB 4; Length 542;

Best Local Similarity 25.3%; Pred. No. 0.84; Mismatches 49; Conservative 19; Indels 60; Gaps 8;
 ; SEQ ID NO: 21396
 ; LENGTH: 542
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-21396

RESULT 9
 US-09-252-991A-19384
 ; Sequence 19384, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/074, 788
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; LENGTH: 1253
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; SEQ ID NO: 30019
 ; US-09-252-991A-30019

Query Match 63 ALVTD--FYDHLESERTQDLFANSTKTVQLKETOQAEYLIGLGRGEYDTEYAAQR-A 117
 Best Local Similarity 22.5%; Pred. No. 3.4%; Mismatches 31; Indels 59; Gaps 11;
 Matches 48; Conservative 31;

Db 181 VILTDRGMFADPRNNPDAQIYEARRDDPLDAVGSGAGALGRGMQTKLRAARLAA 240

Query Match 118 RIGKIHDFVLGQSPDVTLGATTRYTGTLLDALADDVVA- 167
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 49; Conservative 21;

Db 241 RSG----GHTVIVG-----GRIERVLDRLAGERLGLTLPDRSKAAR-KW 283

Query Match 168 VARFLPMKLTFD 181
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 50; Conservative 21;

Db 284 LAGHLOMKGTLVLD 297

RESULT 8
 US-09-252-991A-30019
 ; Sequence 30019, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/074, 788
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; LENGTH: 1253
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; SEQ ID NO: 30019
 ; US-09-252-991A-30019

Query Match 83 ALVTD--FYDHLESERTQDLFANSTKTVQLKETOQAEYLIGLGRGEYDTEYAAQR-A 117
 Best Local Similarity 22.5%; Pred. No. 3.4%; Mismatches 31; Indels 59; Gaps 11;
 Matches 48; Conservative 31;

Db 181 VILTDRGMFADPRNNPDAQIYEARRDDPLDAVGSGAGALGRGMQTKLRAARLAA 240

Query Match 118 RIGKIHDFVLGQSPDVTLGATTRYTGTLLDALADDVVA- 167
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 49; Conservative 21;

Db 241 RSG----GHTVIVG-----GRIERVLDRLAGERLGLTLPDRSKAAR-KW 283

Query Match 168 VARFLPMKLTFD 181
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 50; Conservative 21;

Db 284 LAGHLOMKGTLVLD 297

RESULT 10
 US-09-252-991A-32970
 ; Sequence 32970, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/074, 788
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 32970
 ; LENGTH: 553
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-32970

Query Match 83 ALVTD--FYDHLESERTQDLFANSTKTVQLKETOQAEYLIGLGRGEYDTEYAAQR-A 117
 Best Local Similarity 21.1%; Pred. No. 1.5%; Mismatches 28; Indels 46; Gaps 5;
 Matches 41; Conservative 28;

Db 1024 SGEDDRQHAFELHALOGE--DRLIGVEGVE---DGLIQDQVCA----AFQA 1068

Query Match 61 ADALVTDYDHLESERTQDLF-----ANSTKTVQLKETOQAEYLIGL 104
 Best Local Similarity 21.1%; Pred. No. 1.5%; Mismatches 28; Indels 46; Gaps 5;
 Matches 41; Conservative 28;

Db 1069 AGRLDVFHQFQVEEDVAVAGVHVNRGNRAGAAGRAEHGDEARLVRGIGLRLHAGQA 1128

Query Match 105 RGBNDTTEAQRARIKGTHDVLGLG----PDYDGLGATRYTGTLLDALADDVVA 154
 Best Local Similarity 21.1%; Pred. No. 1.5%; Mismatches 28; Indels 46; Gaps 5;
 Matches 41; Conservative 28;

Db 1129 R-PEEVFGQR----HAYVVGHLGGVEGYLEDGAGV---EYGLLDPG-DHYRA 1177

Query Match 155 DRGEAAAAA-----VDELVARFLPMLKLTFD 181
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 50; Conservative 21;

Db 1178 AQQEDEVVAFPHVARPVGEALAAVVLGLQLVALD 1210

RESULT 9
 US-09-252-991A-19384
 ; Sequence 19384, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/074, 788
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; LENGTH: 1253
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; SEQ ID NO: 19384

Query Match 120 DMRGIDGHALADRIGDEAETAWRLSFTGIDDDTMALAAEQLPFTADALV--TDF 68
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 49; Conservative 21;

Db 129 DARRPADAHPPVQRAF-QATLAQR----QVTTATAANGQRLALAAGSYVAQHQTQY 180

Query Match 69 YDHLESERTQDLFANSTKTVQLKETOQAEYLIGLGRGEYDTEYAAQRARIKGTHDVLG 127
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 50; Conservative 21;

Db 181 ABHQDQEBQASHLDDVPEPLRVERNADLFGLGRGERVYDAAVAGVVAHRL- 239

Query Match 128 LGPDVVLGATRYTGTLLDALAD-----DV-----VADRGEE- 159
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 50; Conservative 21;

Db 240 ---YQAGEAKQFVGAPGLVYDHLVQFHHRROSQAQSLLLGUSVADRAVQF 294

Query Match 160 AAADELVARFLPMLK 177
 Best Local Similarity 24.6%; Pred. No. 1.8%; Mismatches 21; Indels 47; Gaps 8;
 Matches 50; Conservative 21;

Db 295 VVAGGLVALVGRVRQLRL 313

RESULT 11
 US-09-252-991A-32970
 ; Sequence 32970, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196-136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; PRIORITY FILING DATE: 1998-02-18
 ; PRIORITY APPLICATION NUMBER: US 60/074, 788
 ; PRIORITY FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 32970
 ; LENGTH: 553
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-32970

Query Match 75 YERTQDLFANSTKTVQLKETOQAE----YLLGURGERGEYDTEYAAQR-A 121
 Best Local Similarity 21.1%; Pred. No. 1.5%; Mismatches 28; Indels 46; Gaps 5;
 Matches 41; Conservative 28;

Db 197 LHALQQFELDQGLVVLRGDAAQAAHQYEVVGLGQVEAGMVFHFWQRQLGQGG 256

Query Match 20 HALDRI----GLDEAETAWRLSFTGIDDDTMALAAEQLPFTADALVTDYDHL 74
 Best Local Similarity 21.1%; Pred. No. 1.5%; Mismatches 28; Indels 46; Gaps 5;
 Matches 41; Conservative 28;

Db 137 HAPPRPLVSAFFDAPAVGLGPAVEGLAVDEVAQACORPLSFESAEQVRVIRAA 196

Query Match 122 -----IHPL-----GLGPDVYLGAATRYTGTLLDALADDV 153
 Best Local Similarity 21.1%; Pred. No. 1.5%; Mismatches 28; Indels 46; Gaps 5;
 Matches 41; Conservative 28;

Db 257 QVPNGDRYLVIAEGTAAGYGVVVAIIGIVEEEVGPVVEGNAEDRHVHVHVAEARG 316

Query Match 154 ADRGEEAAAVDEL 167
 Best Local Similarity 21.1%; Pred. No. 1.5%; Mismatches 28; Indels 46; Gaps 5;
 Matches 41; Conservative 28;

Db 317 LPASDEFGVALLDD 330

RESULT 11
 US-09-252-991A-18102
 ; Sequence 18102, Application US/09252991A
 ; Sequence 18102, Application US/09252991A

Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 FILE REFERENCE: 107196.136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 PRIOR APPLICATION NUMBER: US 02-18
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS.: 33142
 SEQ ID NO 18102
 LENGTH: 677
 TYPE: PRT
 ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-18102

Query Match 8.6%; Score 80; DB 4; Length 677;
 Best Local Similarity 26.6%; Pred. No. 2;
 Matches 45; Conservative 18; Mismatches 64; Indels 42; Gaps 10;
 Query 6 DPLVTAQVRGIDGHALADRIGDEAFIAWRLSFTGIDDITMAAALAAEQPLFEATADALV 65
 Db 245 DKLVSGDPTPMVADEVDRINAR---SWR-----AFIGDFPLSREDREALI 288
 Query 66 TDFYDHLESERTQDIFANSTKTEQ---LKEIQAEFLL-GLRGEYDTEYYAAQRARI 119
 Db 289 A-----LYEPRDYLAG--KSVEEKETYLAKTTSYDYLKRNVLSETSYKVYFQ---- 334
 Query 120 GKHDVGLGPDVYLGAATRYTGL--LDALADDVADRGEEAAVDE 166
 Db 335 GRSNDFSLAGADA-LPAADA/AAAGFPGSFDALG--.LPQPSSEEAQAMDE 379

RESULT 12
 US-09-540-236-2595
 / Sequence 2595, Application US/09540236
 / Patent No. 6673910
 / GENERAL INFORMATION:
 / APPLICANT: Gary L. Breton et al.
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
 / FILE REFERENCE: 2709-2005-001
 / CURRENT FILING DATE: 2000-04-04
 / NUMBER OF SEQ ID NOS: 3840
 / LENGTH: 438
 / TYPE: PRT
 / ORGANISM: M.catarrhalis
 US-09-540-236-2595

Query Match 8.5%; Score 79; DB 4; Length 438;
 Best Local Similarity 23.8%; Pred. No. 1.4;
 Matches 53; Conservative 27; Mismatches 73; Indels 70; Gaps 10;
 Query 17 IDGHALADRIGDEAFIAWRLSFTGIDDITMAAALAAEQPLFEATADALVTDYDHLSEY 76
 Db 19 IDPQALANALPHF-EYDGTIHFDGINTHILKELYG-TPYAYSORTL-----LENYQ 69
 Query 77 RTQDF-----AST-----KT-----VEQLKETQEYLLG 102
 Db 70 SYTDAFDAIDHQICYAVKANSNLALKTAKAGAFDIVSKGELARVLQADAKKVVTS 129
 Query 103 LGREGYDTYAAQR-----ARIGKTHDVLG-----LGPDVYLGATRYY 141
 Db 130 VGTADDEALKRADLHCFAVESLSBDLTNEAVQRQNQIARISLRNPPDADTHPYIS 189
 Query 142 TGLLDALADDPVADRGEEAAAVDELVARLPLMKLILTEQQI 184
 Db 190 TG---MKGDKFGISHEQAVAY--VYAHHLPLKIVGIDCHI 226

RESULT 14
 US-09-318-352-5599
 / Sequence 559, Application US/09328352
 / Patent No. 6562358
 / GENERAL INFORMATION:
 / APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: GTC 99-13PA
 CURRENT APPLICATION NUMBER: US/09/328,352
 CURRENT FILING DATE: 1999-06-04
 SEQ ID NO: 5599
 LENGTH: 733
 TYPE: PRT
 ORGANISM: Acinetobacter baumannii
 US-09-328-552-5599

Query Match	Match	Score	DB	Length
Best Local Similarity	8.4%	78.5	5	755
Matches	52;	Pred.	No. 3	4;
	Mismatches	32;	Indels	83;
	Gaps	10;		
Qy	9	VTADYRNGIDGHALADR-----IGLDEAETAWRLSTFGIDD-----TMAALA 51		
Db	223	IPAVDNTSSNVAAECKRLTGILG-----LNQGYDKHEHITVNLTVYQAN 273		
Qy	52	APQPLFET----ADALVTDYDFDHLSEYERTQDLFANSTRKE-----90		
Db	274	VERRSAEEAQTLKFLDQLPDLKKQDLDAERQENKFQROQNTVDVTKESELYLQTOSITLE 333		
Qy	91	---QLKETQAEVYLGLRGEGYDTEYAAQR-----ARIGKIHIVLGQGPDV --Y 133		
Db	334	TKKAEELERQAE-----AKYTAEHHPAMREINGQTAINKQLPKDQYQ 388		
Qy	134	LGAY-----TRYTGLILDADDVYDARGE-----EAARAVDELVAFPLMKLIT 179		
Db	389	LQLYREVKTQLTYALLNSYQOLRIAKAGEIGNVRIVTDAVEPIPKPKLQLVLISL 448		
Qy	180	F 180		
Db	449	F 449		

RESULT 15

PCT-US93-07923-3
 Sequence 3, Application PC/TUS93/07923

GENERAL INFORMATION:
 APPLICANT: Morimoto, Chikao
 APPLICANT: Schlossman, Stuart F.
 APPLICANT: Tanaka, Toshiaki
 TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: U.S.A.
 ZIP: 02110-2804

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 COMPUTER: IBM PS/2 Model 50Z or 55SX
 OPERATING SYSTEM: TBM P.C. DOS (Version 3.30)
 SOFTWARE: Wordperfect (Version 5.0)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US93/07923
 FILING DATE: 19930819
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 07/934,162
 FILING DATE: 21-AUG-1992
 APPLICATION NUMBER: 07/832,211
 FILING DATE: 06-FEB-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34,819
 REFERENCE/DOCKET NUMBER: 00530/055002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 755
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 PCT-US93-07923-3

Query	Match	Score	DB	Length
2	SNDNDTIVTADYR-NGIDG---HALADRIGH---DEAETIARLSFTGIDD-----45	8		
Qy	2 SNDNDTIVTADYR-NGIDG---HALADRIGH---DEAETIARLSFTGIDD-----45	8		
Db	558 STENNITVASFDRGRSGYQGDKIMHAIRNRLGTPEVEDOEARQFSKMGFVDNKRIAIWG 617	25	87	
Qy	46 -----TMAALAAEQQLEE---ATA-----DALVTDYDFDHLSEYERTQDLFANSTK 87	3.5		
Db	618 WSYGGYVTSMVIGSGGVFKCGTAVAPVSRMEYYDSYTERMGLPPEDNLDHYRNSTV 677	35	87	
Qy	88 TVEQLKETQAEVYLGLRGEGYDTEYAAQRARIGKIHIVLGQGPDVYILGAYTRYTGLDA 147	35	87	
Db	678 MSRAENPKQVEYLILHGTD-DNVHFQOSAQ1SKA-LVDVGSVDFOAMWYTDDEHGIASS 734	35	87	
Db	735 TA 736	35	87	
Qy	148 LA 149	35	87	
Db	Job time : 9.35593 secs	35	87	
Qy	Search completed: August 10, 2004, 15:30:20	35	87	

Result No.	Score	Query Match	Length	DB ID	Description
SUMMARIES					
1	99	10.6	448	12	US-10-460-524-5
2	94.5	10.1	883	15	Sequence 5, Appli
3	94	10.1	218	14	Sequence 18563, A
4	89.5	9.6	400	14	Sequence 9281, AP
5	88	9.4	505	12	Sequence 13480, A
6	87	9.3	496	12	Sequence 66025, A
7	86	9.2	362	16	Sequence 65364, A
8	85	9.1	2703	12	Sequence 137315, A
9	83.5	8.9	258	14	Sequence 66108, A
10	83.5	8.9	315	12	Sequence 11335, A
11	83.5	8.9	367	12	Sequence 59961, A
12	83.5	8.9	368	12	Sequence 72593, A
13	83.5	8.9	1254	12	Sequence 48921, A
14	83	8.9	384	12	Sequence 57378, A
15	82.5	8.8	403	12	Sequence 48636, A
PREDICTION					
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
ALIGNMENTS					
RESULT 1 US-10-460-524-5					
; Sequence 5, Application US/10460524					
; GENERAL INFORMATION:					
; Publication No. US20040029781A1					
; APPLICANT: Hernan, Ronald A.					
; APPLICANT: Mahigh, Richard J					
; APPLICANT: Brockie, Ian					
; APPLICANT: Jenkins, Elizabeth					
; TITLE OF INVENTION: Affinity Peptides and Method for Purification of Recombinant Proteins					
; FILE REFERENCE: SGM 7047.1					
; CURRENT APPLICATION NUMBER: US 10/460,524					
; PRIORITY APPLICATION NUMBER: US 2003-06-12					
; NUMBER OF SEQ ID NOS: 23					
; SEQ ID NO 5					
; LENGTH: 448					
; TYPE: PRT					
; ORGANISM: Streptococcus					
US-10-460-524-5					
Query Match 10.6; Score 99; DB 12; Length 448;					
Best Local Similarity 25.3%; Pred. No. 0.087;					
Matches 48; Conservative 25; Mismatches 45; Indels 72; Gaps 8;					
QY					
TADYVRNGIDGHALADRGLDEAIAIWNLISFTGIDD-----DTMAIAAE-----53					
41 TPIIRNGGB---LTNLGNSETTLAIRENEESATADLTAAVADTVAAAGAAWEA 97					
QY					
54 ---QPLBEATADAL-----VTDFYDHLESYERTQLFANSTKTYVQLKETQAEYLIG 102					
98 AAADAKAKADAKRPEKKIGSDYKRN-----INNAKTYEGIKDQLQAQV-- 145					
103 LGREGYDTEYAAQRARIGKIHDLVGLGPDVYLVGATYYTGLLDAADDVVAADRGEAAA 162					
Db					
146 -----ESARKARISEATD-----GLSDFLKSKQTPA--EDTVK 175					

APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 CURRENT APPLICATION NUMBER: US/10/282,122A
 PRIORITY FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 FILE REFERENCE: ELITRA_034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 REMAINING PRIOR APPLICATION DATA REMOVED - SEE FILE WRAPPER OR PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 66025 LENGTH: 505 TYPE: PRT
 ORGANISM: Neisseria meningitidis
 US-10-282-122A-66025

Query Match 9.4%; Score 88; DB 12; Length 505;
 Best Local Similarity 25.7%; Pred. No. 1.5;
 Matches 46; Conservative 24; Mismatches 69; Indels 40; Gaps 7;

Qy 6 DTLVTAADVNGIDGHALADRIGLDEAEIANRSLP---TGIDDDTMALAAEQPLF----
 Db 316 DQLALIIGRGQNVRASDLTG----WQLNIMTSAEDERNAEDAIIRLFMDHN 368

Qy 58 --EATADALYTDFYDHL--SYERTQDLFANSTKTVBQLKETQAELYLLGLGRGEYDTEYA 113
 Db 369 VDEETADVLQEGFATLEEVAYVPAELLA----IEGFDEETVDMLRNRDALTMAI 423

Qy 114 AQRARIKGKHHDVL---GLOPDVYLGAYTRYTYGLLDALADDVVYDRGEAAAAYDELV 168
 Db 424 AAEEKIGEVSDDMRNLEGVDAMLL----SLEAGITTRDDLAEALAVDELI 470

RESULT 7
 US-10-437-963-137315
 Sequence 137315, Application US/10437963
 GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovacic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21-(5322) B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 20496
 ; SEQ ID NO: 137315 LENGTH: 362
 ; TYPE: PRT
 ; ORGANISM: Oryza sativa

RESULT 6
 US-10-282-122A-65364
 Sequence 65364, Application US/10282122A
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cherry
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_3880C.1.pep
US-10-437-963-137315

Query Match 9.2%; Score 86; DB 16; Length 362;
Best Local Similarity 26.4%; Pred. No. 1.6;
Matches 51; Conservative 20; Mismatches 78; Indels 44; Gaps 9;

Qy 8 LVTADYNGIDGHALADRIGLDEAEIAWR---LSFT-GIDDDTMAAHLAREQPLFEATA 61
Db 26 IIAADESTGTGKRLIAS-IGVNEVEENRBALELLTAPGALDCLSGVILFEETLYQSTR 84

Qy 62 DALVTDFYDHLESY-----ERTQDLFANSTKTVQEQLKETOQAEYLGLGRGEYDTEY 112
Db 85 DG---TPPVDVLAAGVLAGIVKIDKGTVELAGTDRETTQGD-----GLGERCRY 133

Qy 113 AAQRARIGKIHDLVGLG-----PDVYLGAYTR-----YTGIDLALADDVYAD-- 155
Db 134 YAAGRFAKWRAVLSTGRASSRPSQLAVDANQGLARYAIICQENGVLPIVEPEILDGE 193

Qy 156 RQEAAIAANDLEY 166
Db 194 HGIERCAEVTERY 206

RESULT 8
US-10-282-122A-66108
Sequence 66108, Application US/10282122A
Publication No. US2004029123A1

GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cherry
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA_034A

CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 78614
LENGTH: 2703
TYPE: PRT
ORGANISM: Neisseria meningitidis

SOFTWARE: Parent In version 3.1
SEQ ID NO 66108

RESULT 9
US-10-156-761-11335
Sequence 11335, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADYOUSHI
APPLICANT: SAKAKI, YOSHIOUKI
APPLICANT: HATORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11335

Query Match 9.1%; Score 85; DB 12; Length 2703;
Best Local Similarity 23.5%; Pred. No. 3.2;
Matches 52; Conservative 26; Mismatches 71; Indels 72; Gaps 9;

Qy 3 NDNTLVTDYVRNGIDGHAL-----ADBLGDAEIAWLTSFTGIDDDT-----46
Db 2153 SDGS-KNSTDGS-KNSTDS-QNTNHTIDAGLARTGRATEKAFARI-YTGIDDETAODHSGLH 2210

Qy 47 -----MAALAAEOPPLFE-----ATADALVTDYDHLESYETQDIFANSTKTVEQ 91
Db 2211 KNSFDKDAVAKINLOREVTEKFGRMNAAQAVAAVADKLQNTSYQEA--RTLLEAE 2267

Qy 92 LKETQFLLGJLGRGEYDTEYAAQRARIGKIHDLV-----GLGPDVYLGAYT 138
Db 2268 LQNTDSE-----AEKAFAFRASLGQNAYLAENQSRYDTWKEGGIGRSILHGAAG 2316

Qy 139 RYYTGLLDAL-----ADVVAADRGEAAAAYDEL 167
Db 2317 GUTTGSUGGILLAGGGTSLAAPYLDKAENLCPAGKAVNAL 2357

RESULT 9
US-10-156-761-11335
Sequence 11335, Application US/10156761
Publication No. US20030119018A1

GENERAL INFORMATION:
APPLICANT: OMURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADYOUSHI
APPLICANT: SAKAKI, YOSHIOUKI
APPLICANT: HATORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 11335

Query Match 8.9%; Score 83.5; DB 14; Length 258;
Best Local Similarity 24.1%; Pred. No. 1.9;
Matches 49; Conservative 21; Mismatches 78; Indels 55; Gaps 9;

Qy 8 LYTDYVRNGIDGHALADRIGLDEAEIAWRSLTSFTGIDDDTMAALAEBQPLFEATADALVTD 67
Db 33 VVADVLDD-QGPALAKETGARYVHLD-----VGREDDWQAAY-----TVAKD 74

Qy 68 FYDHESYE-----RTQDLFANSTKTVEQLKE-TQAEYLLGIIGRGEYDTEAAQRARIG 120
Db 75 AYGHDGLVNNAAGLRLRENDFLVGTPLAEPQQVQNVQGVFLGKTVAPBIE-AAGGGTTV 133

Qy 121 KIHDLVGLGPDVYLGAYT-----RYYTGLLD-ALADDVVA 154
Db 134 NTASYAGLTGMAYGATKHAIGLTRVAALELAACKIRVNAVCPCGAIDTAMSNPSQL 193

Qy 155 DRG---BEAAAAYDELVARFLPM 174
Db 194 DPGDPEETARALSBLYXGRLVPL 216

RESULT 10
US-10-282-122A-59961

Sequence 59961, Application US/10288122A
 Publication No. US2004029129A1
 GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zamudio, Carlos ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Malone, Cheryl ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Haselbeck, Robert ; APPLICANT: Soren, Steven E.
 ; APPLICANT: Ohlsen, Kari ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Zyskind, Judith ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wall, Daniel ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; APPLICANT: Trawick, John ; FILE REFERENCE: 38-21(53313) B
 ; APPLICANT: Carr, Grant ; CURRENT APPLICATION NUMBER: US/10/425, 114
 ; APPLICANT: Yamamoto, Robert ; CURRENT FILING DATE: 2003-04-28
 ; APPLICANT: Forsyrb, R. ; NUMBER OF SEQ ID NOS: 73128
 ; APPLICANT: Xu, H. ; SEQ ID NO: 72593
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms ; LENGTH: 367
 FILE REFERENCE: ELITRA 034A ; Score: 83.5; DB 12; Length 367;
 CURRENT APPLICATION NUMBER: US/10/282,122A ; Best Local Similarity: 25.3%; Pred. No. 3;
 CURRENT FILING DATE: 2003-02-20 ; Matches: 47; Conservative: 24; Mismatches: 68; Indels: 47; Gaps: 10;
 PRIOR APPLICATION NUMBER: 60/191,078 ; OTHER INFORMATION: Clone ID: CC-ZMPO14820017E02_FLI.pep
 PRIOR FILING DATE: 2000-03-21 ; US-10-425-114-72593
 PRIOR APPLICATION NUMBER: 60/206,848 ; Query Match: 8.9%; Score: 83.5; DB 12; Length 367;
 PRIOR FILING DATE: 2000-05-23 ; Best Local Similarity: 25.3%; Pred. No. 3;
 PRIOR APPLICATION NUMBER: 60/207,727 ; Matches: 47; Conservative: 24; Mismatches: 68; Indels: 47; Gaps: 10;
 PRIOR APPLICATION NUMBER: 60/230,335 ; OTHER INFORMATION: Zea mays ; FEATURE:
 PRIOR FILING DATE: 2000-09-06 ; Qy 16 GIDGHALADRIGLDEAEIAWRLSFTGIDDDTMAALAEQPLFEATDALVTDFYDHESY 75
 PRIOR FILING DATE: 2000-09-09 ; Db 152 GVDGR---RVLGLDEBAP---DADAVAAAGQRDHILD--WEALLYVELVDELGH- 196
 PRIOR APPLICATION NUMBER: 60/242,578 ; Qy 76 ERTQDIFANSTKTV----BQLKETQAELYLLGLR----GEYDTEYAAQ----R 116
 PRIOR FILING DATE: 2000-10-23 ; Db 197 --RDYTRARQYGVGLAVGHERVAPPGQHRYLGPQLHRVAGAHRRDVRARKHVRALLQ 253
 PRIOR APPLICATION NUMBER: 60/253,,625 ; Qy 117 ARIKGHDVILGLGPDV---YLGAATRY---YTGLLDALADDVVAADRGEEAAAAYDELVA 169
 PRIOR FILING DATE: 2000-11-27 ; Db 254 IRLDLIDHVVALHKGDGAGLGLGGVLRHRVDQDGRVAACIEAVVDEAEAGGGAG--V 310
 PRIOR APPLICATION NUMBER: 60/257,931 ; Qy 170 RFLPML 175
 PRIOR FILING DATE: 2000-12-22 ; Db 311 SFEPUL 316
 PRIOR APPLICATION NUMBER: 60/267,636 ;
 PRIOR FILING DATE: 2001-02-09 ;
 PRIOR APPLICATION NUMBER: 60/269,,308 ;
 PRIOR FILING DATE: 2001-02-16 ;
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614 ;
 SOFTWARE: PatentIn version 3.1 ;
 SEQ ID NO: 59961 ;
 LENGTH: 315 ;
 TYPE: PRT ;
 ORGANISM: Klebsiella pneumoniae ;
 US-10-282-122A-59961 ;
 Query Match: 8.9%; Score: 83.5; DB 12; Length 315;
 Best Local Similarity: 25.4%; Pred. No. 2.4; Mismatches: 51; Conservative: 17; Mismatches: 61; Indels: 61; Gaps: 9;
 Matches: 51; Conservative: 17; Mismatches: 72; Indels: 61; Gaps: 9;
 Qy 14 RNGIDGHALADRIGLDEAEIA-WRLSFTGIDDDTMAALAEQPLF----EATADALIV 65
 Db 116 RRNIDGVVLFGLFTGIDEAMAPWR-----DTLVLMARDAPSFAASYCDDGAIILM 167
 Qy 66 TDFYDH-----LESVERTQDIFANSTKTVQAKETQAELYLLGLRGEYDTE 111 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 Db 168 QRLYDRGRHHSISFLGVPHSDVTTGERRHLAYLAFCKK-HRLTPTAAFLGLGMFOG-YDTV 225 ; FILE REFERENCE: 38-21(53313) B
 Qy 112 YAQRARIGKI---HDVGLGPDVYLGAYTRYTGGLD-----148 ; CURRENT APPLICATION NUMBER: US/10/425, 114
 Db 226 ASVLTAEISALVCAATDTLAGSKYLOQQR----DALQLASVGSTPLMKFLPEILT 279 ; CURRENT FILING DATE: 2003-04-28
 Qy 149 ADDVVAADRGEAAAAYDELVA 169 ; NUMBER OF SEQ ID NOS: 73128
 Db 280 VDPGYAESGRRAQRQLIBQIA 300 ; LENGTH: 368;
 Qy 16 GIDGHALADRIGLDEAEIAWRLSFTGIDDDTMAALAEQPLFEATDALVTDFYDHESY 75 ; OTHER INFORMATION: Clone ID: CC-ZMPO148032B07_FLI.pep
 RESULT 11 ;
 US-10-425-114-72593 ;
 Sequence 72593, Application US/10425114 ;
 Publication No. US2004003488BA1

42 IDDDTMAALAAEOPFEATADALYTFDHDLESYERTQDFANSTKTVQELKETQAEYLL 101
 Qy 717 AGPOYQA--AADSAAKEGTPDAL--DEFPGDGYKAR--YLDQWQQAVELDT----- 764
 Db 76 ERTQDLFANSTKTIV----TQLKETQAEYLLGLGR----GEVDTEYAAQ----R 116
 Qy 197 --RDVTRQVGVLAVGHERVAPQGHVRVLPAROLHRVAGAIRDVRHVRALIQLQ 253
 Db 117 ARIGKTHDVILGLGPDY----YLGATRY----YTGLIDALADDVYADRGEAAAAYDELVA 169
 Qy 254 LRLDLHDYVLAHLGDVGAQLLGGIVRHRVQDGRVAAICEAVVDEEABGGGAG--V 310
 Db 170 RFLPML 175
 Qy 311 SFEPLI 316

RESULT 14
 US-10-425-114-48636
 Sequence 48636, Application US/10425114
 Publication No. US20040103488A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA_034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See file wrapper or PALM.
 NUMBER OF SEQ ID Nos: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 53778
 LENGTH: 1254
 TYPE: PRT
 ORGANISM: Corynebacterium diphtheriae
 US-10-28-122A-53778

Query Match 8.9%; Score 83.5; DB 12; Length 1254;
 Best Local Similarity 25.0%; Pred. No. 16; Mismatches 82; Indels 51; Gaps 10;
 Matches 52; Conservative 23;

4 DNDTYL-----TADVNRGIDGHALADRIGLE-----AEIAWRLSFG 41
 Qy 658 DNLNLYNHWAFLDPNEDLNRNADEACYDITDEFVNITAPQRNLPGVYKAMELR-DG 716

RESULT 15
 US-10-282-122A-61978
 Sequence 61978, Application US/10282122A
 Publication No. US200402129A1
 GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Kari
 APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA_034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO: 61978
 LENGTH: 403
 TYPE: PRT
 ORGANISM: Mycobacterium avium
 US-10-282-122A-61978

Query Match 8.8%; Score 82.5; DB 12; Length 403;
 Best Local Similarity 25.4%; Pred. No. 4.4;
 Matches 36; Conservative 20; Mismatches 65; Indels 21; Gaps 6;
 Qy 42 IDDDTMAAAEQQPLFFATADALVTDYDHL--ESYERTQDLFANSTKTVQLKETQAEY 99
 Db 25 ISDSVLDALLAQDPRSRVAELVLVIGQVHVGIVTTTAKEEAFADTTNTVR-----ER 77
 Qy 100 LLGLGRGEYDTEYYAAQRARIGKTHDVLGLG--PDVYLGLAYTRYT--GILDAIADDVV 153
 Db 78 ILDIGDSSDKGFDSGCGN----IGIGAQSPDIAQGVTDAHETRVEGAADPLDAQGA 132
 Qy 154 ADRGEFAAAAVDELVRF-LPM 174
 Db 133 GDQGLMFGYZAINDTPERMPLPI 154

Search completed: August 10, 2004, 15:42:28
 Job time : 27.7312 secs

Blank

Copyright (c) 1993 - 2004 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: August 10, 2004, 15:22:45 ; Search time 8.01937 Seconds
 (without alignments)
 2207.061 Million cell updates/sec

Title: US-09-455-978B-77

Perfect score: 933

Sequence: 1 MSNDNTLVADVRNGIDGH.....DELVARFLPMLKLTDFDQI 184

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:

283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:
 1: Pir1:
 2: Pir2:
 3: Pir3:
 4: Pir4:
 *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	933	100.0	489	2 E84304	Htr10 transducer [imported] - Halobacterium sp. NRC-1
2	933	100.0	489	2 T44978	Htr10 transducer [imported] - Halobacterium sp. NRC-1
3	148	15.9	439	2 A83713	C;Species: Halobacterium sp. NRC-1
4	142	15.2	432	2 C69832	C;Accession: E84304
5	132	14.1	537	2 A87302	C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
6	117.5	12.6	499	2 A97482	R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonski, K.H.; Alam, M.; Freitas, T.; Jung, K.H.; Alam, M.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A;Title: Genome sequence of Halobacterium species NRC-1.
7	117.5	12.6	499	2 AB2703	A;Reference number: A84160; PMID:20504483; PMID:11016930
8	105.5	11.3	555	2 D87536	A;Accession: E84304
9	99	10.6	448	2 A24496	A;Status: Preliminary
10	98	10.5	593	2 S00128	A;Residues: 1-489 <STOP>
11	97	10.4	218	2 T35174	A;Cross-references: GB:AE004437; PIDN:AG19801.1; GSPDB:GN00138
12	94.5	10.1	883	2 A84210	C;Genetic:
13	94	10.1	1090	2 T00533	C;Gene: hr10
14	93.5	10.0	2048	2 C84609	C;Superfamily: Halobacterium salinarum transducer protein htr1
15	93	10.0	881	2 AD2580	Query Match 100.0%; Score 933; DB 2; Length 489;
16	93	10.0	881	2 C97362	Best Local Similarity 100.0%; Pred. 1.9e-70; Mismatches 0; Indels 0; Gaps 0;
17	92.5	9.9	503	2 AH3535	QY 1 MSNDNTLVADVRNGIDGHALADRGLEDEIAWRLSFPGIDDMMALAEEQLFEAT 60
18	89	9.5	955	2 A47334	Db 1 MSNDNTLVADVRNGIDGHALADRGLEDEIAWRLSFPGIDDMMALAEEQLFEAT 60
19	88	9.4	500	2 BB1060	QY 61 ADALVTDYDHLESVERTODLNFANSTKTYEQLEAQYLLGIGREYDTEYAAQRARTG 120
20	88	9.4	505	2 BB1816	Db 61 ADALVTDYDHLESVERTODLNFANSTKTYEQLEAQYLLGIGREYDTEYAAQRARTG 120
21	88	9.4	775	2 T45238	QY 121 KIHDLVLGIGLPDVYLGAATRYTGTGLDAADDVYADRGREAAAAYAVDELYAREFLDMKLITP 180
22	87	9.3	469	2 BB7094	Db 121 KIHDLVLGIGLPDVYLGAATRYTGTGLDAADDVYADRGREAAAAYAVDELYAREFLDMKLITP 180
23	86.5	9.3	157	2 F84349	RESULT 2
24	85	9.1	2703	2 HB1193	T44978
25	84.5	9.1	266	2 C84365	transducer protein hemA [validated] - Halobacterium salinarum
26	84	9.0	491	2 F96022	C;Alternates: methyl-accepting taxis protein htB; transducer protein htB; transducer protein hemA [validated]
27	83.5	8.9	712	2 T20552	C;Species: Halobacterium salinarum
28	82.5	8.8	381	2 S28115	C;Accession: T44978
29	82	8.8	331	2 S01964	R;Zhang, W.; Brooun, A.; McCandless, J.; Banda, P.; Alam, M.

Proc. Natl. Acad. Sci. U.S.A. 93, 4649-4654, 1996

Run on:	August 10, 2004, 15:22:45 ; Search time 8.01937 Seconds
Title:	US-09-455-978B-77
Perfect score:	933
Sequence:	1 MSNDNTLVADVRNGIDGH.....DELVARFLPMLKLTDFDQI 184
Scoring table:	BLOSUM62
Gapop:	10.0
Gapext:	0.5
Scored:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3: Pir3: 4: Pir4: *
Scoring table:	BLOSUM62
Gapop:	10.0 , Gapext 0.5
Scanned:	283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters:	283366
Minimum DB seq length:	0
Maximum DB seq length:	2000000000
Post-processing: Minimum Match 0%	Maximum Match 100%
Database :	PIR 78: 1: Pir1: 2: Pir2: 3:

A;Title: Signal transduction in the archaeon Halobacterium salinarium is processed through a two-component system. A;Reference number: Z22804; MUID:96209786; PMID:8643458
A;Accession: F44978
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-489 <ZHA>
A;Cross-references: EMBL:U75436; NID:gi1654420; PID:AAB17881.1; PMID:91654421
A;Experimental source: strain Fix15
A;Gene: hemAT; htbp1; htB
C;Function: involved in aerotactic signal transduction; involved in oxygen sensing; transduction protein
C;Description: involved in aerotactic signal transduction; involved in oxygen sensing; transduction protein
C;Superfamily: Halobacterium salinarum transducer protein htrI
C;Keywords: heme; methylated amino acid; signal transduction

Query Match Score 933; DB 2; Length 489;
Best Local Similarity 100.0%; Pred. No. 1.9e-70;
Matches 184; Conservative 0; Gaps 0; Indels 0;

Db 1 MSNDNTLTVTADVRNGDGHALADRIGLDDEAEIANTVLSFTGDDDTMAALAAEQPIFEAT 60
Db 1 MSNDNTLTVTADVRNGDGHALADRIGLDDEAEIANTVLSFTGDDDTMAALAAEQPIFEAT 60

Qy 61 ADALVTDYDHLESYERTQDFANSTKIVQEQLKETQAEYLGLGREYDTYYAQARIG 120
Db 61 ADALVTDYDHLESYERTQDFANSTKIVQEQLKETQAEYLGLGREYDTYYAQARIG 120

Qy 121 KIHDVIGLGPDVYLGAATRYTGLLDAADDVVAADDELVARFLPMLKLTF 180
Db 121 KIHDVIGLGPDVYLGAATRYTGLLDAADDVVAADDELVARFLPMLKLTF 180

Qy 181 DQOI 184
Db 181 DQOI 184

RESULT 3
A83713 accepting chemotaxis protein BH0505 [imported] - Bacillus halodurans (strain C-12)
C;Species: Bacillus halodurans
C;Accession: A83713
C;Genetics: BH0505
A;Title: Complete genome sequence of the alkaliophilic bacterium *Bacillus* halodurans and its pathogenicity. A;Reference number: A83650; PMID:20512582
A;Accession: A83713
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-439 <STD>
A;Cross-references: GB:AP001508; GB:BA000004; NID:91017890; PID:BAB04224.1; GSPPDB:GN000
A;Experimental source: strain C-125

Query Match Score 148; DB 2; Length 439;
Best Local Similarity 25.2%; Pred. No. 8e-05; Gaps 3; Indels 10; Mismatches 71;

Db 30 EAETAWRSLFTGDDDTMAALAAEQPIFEATADALVTDYDHLESYERTQDLFANSKTV 89
Db 33 ESELSAQKRMHITLDDDKRMKAQLOPMEYLAQAYSNIKOPNINELL-EHTSV 91

Qy 90 QLKETQAEYLGLGREYDTYYAQARIGKHDVIGLGPDVYLGAATRYTGLLDA 149
Db 92 ERLKETLKQHILEMENGEDQAFQKRLQIAQAH-----VRIGLQTKWVSAFOQLT 143

Qy 150 DDVYADRGEEAAAAYDELVARFLPMLKLTFDQOI 184
Db 144 DSL1-QLEOHLSQSPSDIVLATRSILKLNLEQQL 177

RESULT 4
C69832 methyl-accepting chemotaxis protein homolog YhfV - *Bacillus subtilis*
C;Species: *Bacillus subtilis*
C;Accession: C69832 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho A.; Ehrlich, S.D.; Emmerson, P.T.; Entwistle, K.D.; Errington, J.; Fabre, C.; Ferrari, E.; Ichel, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Huilo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapudis, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, V.; Liu, H.; Masuda, S.; Maezel, Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, T.N.; Portebeille Rieger, M.; Rivolta, C.; Rocha, B.; Roche, B.; Rose, M.; Sekiguchi, J.; Sescano, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Tognoni, A.; Uchiyama, T.; Winters, P.; Yipat, A.; Yoshikawa, H.R.; Yoshimoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshioka, H.R.; Zumstein, E.; Yoshihiko, H.; Danchin, A.;Title: The complete genome of the Gram-positive bacterium *Bacillus subtilis*. A;Reference number: A69580; PMID:9804433; NID:92633260; PID:912633374
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Accession: C69832
A;Experimental source: strain 1868
A;Gene: yhfV
C;Superfamily: Pyrococcus horikoshii probable methyl-accepting chemotaxis protein; trans

Query Match Score 142; DB 2; Length 432;
Best Local Similarity 25.3%; Pred. No. 0.0035%; Gaps 2; Indels 10; Mismatches 76;

Db 31 ABIAWRLSFTGIDDDTMAALAAEQPIFEATADALVTDYDHLESYERTQDLFANSKTV 90
Db 32 ADVKQLMVRQGDAELVLYEQLPLQIQTENIIVIDAFYKNLDHESSLMIDII-NDHSSVD 90

Qy 31 ABIAWRLSFTGIDDDTMAALAAEQPIFEATADALVTDYDHLESYERTQDLFANSKTV 90
Db 32 ADVKQLMVRQGDAELVLYEQLPLQIQTENIIVIDAFYKNLDHESSLMIDII-NDHSSVD 90

Qy 91 QLKETQAEYLGLGREYDTYYAQARIGKHDVIGLGPDVYLGAATRYTGLLDA 150
Db 91 RLKQTLKRHIQMFMAGVTDIEFKRNTRIASHLRIGLKPQYMGAFQELLSMSMDIY-- 148

Qy 151 DVADGRGEAAAADVDELVARFLPMLKLTFDQOI 184
Db 149 -----EASITNQQELKAIKATKLNLEQQL 175

RESULT 5
C87302 methyl-accepting chemotaxis protein McpB [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Accession: C87302 #sequence_revision 20-Apr-2001 #text_change 17-May-2002
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolonowicz, M.J.; Birmajera, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venturini, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 41336-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87302
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-537 <STD>
A;Cross-references: GB:AE005673; NID:913421593; PID:AAK22415.1; GSPPDB:GN00148
A;Gene: CC0428
C;Superfamily: methyl-accepting chemotaxis protein mcpA
Query Match Score 132; DB 2; Length 537;
Best Local Similarity 27.2%; Pred. No. 0.0023%; Gaps 12; Mismatches 72;

Db 41; Conservative 41; Gaps 41; Mismatches 26;

Qy 33 IAWRLSFTGIDDDTMALAAEOPLFATADALVTDFYDHLSEYERTODLFANSTKTVEQL 92
 Db 11 IGERTAMGIDDKARSALRDLRPTVIREIGKALDNFVGKVRATPTEKFSDDRHMAAS 70
 C;Species: Agrobacterium tumefaciens
 C;Accession: AB2203
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourollo, B.; Goldman, R.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2223-2228, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens: 1.499 <KUR>
 A;Accession: A97485
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-499
 A;Cross-references: GB:AE007869; PIDN:AAK86834.1; PMID:21608551; PMID:11743194
 A;Genetics:
 A;Gene: AGR_C_1888
 A;Map position: circular chromosome
 A;Map position: circular chromosome

Query Match 12.6%; Score 117.5; DB 2; Length 499;
 Best Local Similarity 26.5%; Pred. No. 0.034; Mismatches 23; Indels 7; Gaps 3;
 Matches 35; Conservative

Qy 19 GHALADRIGLDEAIIAWRLSFTGIDDDTMALAAEOPLFATADALVTDFYDHLSEYERT 78
 Db 3 GOAKTR-QLDL---RINFLGLGHGERQNLSMDKGVTISLDASIDRFTKVRAPET 56
 C;Species: Streptococcus sp. (group G)
 C;Accession: A24496
 C;Date: 17-Sep-1987 #sequence_revision 17-Sep-1987
 R;Fahnestock, S.R.; Alexander, P.; Nagle, J.; Filipua, D.
 J. Bacteriol. 167, 970-880, 1986
 A;Title: Gene for an immunoglobulin-binding protein from a group G streptococcus
 A;Reference number: A24496
 A;Accession: A24496
 A;Molecule type: DNA
 A;Residues: 1-448 <KUR>
 A;Cross-references: GB:AE008888; PIDN:AAL42440.1; PMID:917739417; GSFDB:GN00186
 R;Stoebring, U.; Bjoerck, L.; Kastern, W.
 J. Biol. Chem. 266, 399-405, 1991
 A;Title: Streptococcal protein G. Gene structure and protein binding properties.

RESULT 6
 methyl-accepting chemotaxis protein mcpV (AF312877) [imported] - Agrobacterium tumefaciens
 C;Species: Agrobacterium tumefaciens
 C;Accession: A97485
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 20-Apr-2001
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ourollo, B.; Goldman, R.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2223-2228, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens: 1.499 <KUR>
 A;Accession: A97485
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-499
 A;Cross-references: GB:AE007869; PIDN:AAK86834.1; PMID:915156046; GSFDB:GN00169
 A;Genetics:
 A;Gene: AGR_C_1888
 A;Map position: circular chromosome
 A;Map position: circular chromosome

Query Match 12.6%; Score 117.5; DB 2; Length 499;
 Best Local Similarity 26.5%; Pred. No. 0.034; Mismatches 23; Indels 7; Gaps 3;
 Matches 35; Conservative

Qy 19 GHALADRIGLDEAIIAWRLSFTGIDDDTMALAAEOPLFATADALVTDFYDHLSEYERT 78
 Db 3 GOAKTR-QLDL---RINFLGLGHGERQNLSMDKGVTISLDASIDRFTKVRAPET 56
 C;Species: Streptococcus sp. (group G)
 C;Accession: A24496
 C;Date: 17-Sep-1987 #sequence_revision 17-Sep-1987
 R;Fahnestock, S.R.; Alexander, P.; Nagle, J.; Filipua, D.
 J. Bacteriol. 167, 970-880, 1986
 A;Title: Gene for an immunoglobulin-binding protein from a group G streptococcus
 A;Reference number: A24496
 A;Accession: A24496
 A;Molecule type: DNA
 A;Residues: 1-448 <FAH>
 A;Cross-references: GB:AE008888; PIDN:AAL42440.1; PMID:917739417; GSFDB:GN00186
 R;Stoebring, U.; Bjoerck, L.; Kastern, W.
 J. Biol. Chem. 266, 399-405, 1991
 A;Title: Streptococcal protein G. Gene structure and protein binding properties.

RESULT 8
 methyl-accepting chemotaxis protein McpV [imported] - Caulobacter crescentus
 C;Species: Caulobacter crescentus
 C;Accession: D87536
 B.;Iaub, M.T.; Debey, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haff, D.H.; Kolon, N.;J.;Bromolaeva, M.;White, O.;Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A;Title: Complete Genome Sequence of Caulobacter crescentus
 A;Reference number: A87249; PMID:2117368; PMID:11259647
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-55 <STO>
 A;Cross-references: GB:AE005673; NID:913423838; PIDN:AAK24288.1; GSFDB:GN00148
 A;Genetics:
 A;Gene: CC2317

Query Match 11.3%; Score 105.5; DB 2; Length 555;
 Best Local Similarity 24.7%; Pred. No. 0.39; Mismatches 28; Indels 17; Gaps 5;
 Matches 40; Conservative

Qy 31 AEIAWRLSFTGIDDDTMALAAEOPLFATADALVTDFYDHLSEYERTQDFANSTKTVE 90
 Db 46 AKLDORMAFRFDERSAHRAIRKEVIDEIGAAGQFSQVRLLPDTVKFRDGH-MA 104
 C;Species: Streptococcus sp.
 C;Accession: A24496; A39041
 C;Date: 17-Sep-1987 #sequence_revision 17-Sep-1987
 R;Fahnestock, S.R.; Alexander, P.; Nagle, J.; Filipua, D.
 J. Bacteriol. 167, 970-880, 1986
 A;Title: Gene for an immunoglobulin-binding protein from a group G streptococcus
 A;Reference number: A24496
 A;Accession: A24496
 A;Molecule type: DNA
 A;Residues: 1-448 <FAH>
 A;Cross-references: GB:AE008888; PIDN:AAL42440.1; PMID:917739417; GSFDB:GN00186
 R;Stoebring, U.; Bjoerck, L.; Kastern, W.
 J. Biol. Chem. 266, 399-405, 1991
 A;Title: Streptococcal protein G. Gene structure and protein binding properties.

RESULT 9
 methyl-accepting chemotaxis protein mcpV [imported] - Agrobacterium tumefaciens (strain C58)
 C;Species: Agrobacterium tumefaciens
 C;Accession: AB2203
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Erae, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan; Karp, P.; Romero, P.; Zhang, S.; Science 294, 2317-2323, 2001
 A;Author: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, M.; Gordon-Kamm, C58.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Accession: AB2203
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-499 <KUR>
 A;Cross-references: GB:AE008888; PIDN:AAL42440.1; PMID:917739417; GSFDB:GN00186
 R;Stoebring, U.; Bjoerck, L.; Kastern, W.
 J. Biol. Chem. 266, 399-405, 1991
 A;Title: Streptococcal protein G. Gene structure and protein binding properties.

A; Reference number: A39041; MUID:91093154; PMID:1985908

A; Accession: A39041

A; Status: preliminary

A; Molecule type: protein

A; Residues: 34-42; N⁺,45-48;62-76;186-200 <S01>

C; Genetics:

A; Gene: B9g

C; Superfamily: M5 protein

C; Keywords: transmembrane protein

Query Match 10.6%; Score 99; DB 2; Length 448;

Best Local Similarity 25.3%; Pred. No. 1;

Matches 48; Conservative 25; Mismatches 45; Indels 72; Gaps 8;

Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53

Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97

Qy 54 ---OPLFEATADL----VTDPFYDHLSEYERTODLFANSTKTEVOLKEETOAYLLG 102

Db 98 AAAADALAKARAKADALREFNKGQSDFKNL----INNAKTVGIRDQLQNV-- 145

Qy 103 LGRGEYDTEYAAQRARIGKIHDLGLCPDVLYGATRYTGTLLADDFVVAADRGEAAA 162

Db 146 -----ESARKARISEATD----GLSDFLRKSQTPA---EDTVK 175

Qy 163 AVDELVARFL 172

Db 176 SIELABAKVL 185

RESULT 10

S00128 protein G precursor - Streptococcus sp. (Streptococcus G148)

N Alternate names: albumin-binding protein; cell wall-bound protein

C; Species: Streptococcus sp.

A; Variety: Streptococcus G148

C; Date: 30-Jun-1989 #text_change 30-Jun-1989 #text_change 19-May-2000

C; Accession: S00128; A27604; T26314

R; Eliasson, A.; Eliasson, M.; Guss, B.; Nilsson, B.; Hellman, U.; Lindberg, M.; Uhlen, M.

Bur. J. Biochem. 168, 319-324, 1987

A; Title: Structure and evolution of the repetitive gene encoding streptococcal protein G

A; Reference number: S00128; MUID:88029445; PMID:3665928

A; Accession: S00128

A; Molecule type: protein

A; Residues: 1-593 <OLS>

A; Cross-references: EMBL:X06173; PIDN:CAA29540.1; PID:947085

A; Note: The source is designated as Streptococcus G148

A; Note: Part of this sequence, including the amino end of the mature protein, was confirmed by sequencing.

R; Sjoerbring, U.; Falkenberg, C.; Nielsen, B.; Akerstroem, B.; Bjoerck, L.

J. Immunol. 140, 1595-1599, 1988

A; Title: Isolation and characterization of a 14-kDa albumin-binding fragment of streptococcal protein G

A; Accession: A27604; MUID:88154455; PMID:2831269

A; Molecule type: protein

A; Residues: 62-101 <S01>

P; Guss, B.; Eliasson, M.; Olsson, A.; Uhlen, M.; Frej, A.K.; Jornvall, H.; Flock, J.I.

EMBO J. 5, 1567-1575, 1986

A; Title: Structure of the IgG-binding regions of streptococcal protein G.

A; Reference number: A26314; MUID:86300657; PMID:3017704

A; Accession: A26314

A; Molecule type: DNA

A; Residues: 114-593 <SUS>

A; Cross-references: GB:X04015; PID:g47071; PID:g47072

C; Description: it is part of the cell wall structure of group G streptococci and is cova-

C; Superfamily: M5 protein

C; Keywords: duplication; membrane protein

F; 1-33;/Domain: signal sequence #status predicted <SIG>

F; 34-593;/Product: protein G #status experimental

F; 34-116;/Domain: AB duby-duplication <DUPI>

F; 117-140;/Region: A repeat

F;141-191/Region: B	B								
F;192-215/Region: A repeat									
F;216-266/Region: B									
F;267-290/Region: A repeat									
F;303-497/Domain: IgG binding <IGB>									
F;358-377/Region: C repeat									
F;373-427/Region: D									
F;428-442/Region: C repeat									
F;443-497/Domain: proline-rich <PRO>									
F;568-593/Domain: carboxyl-terminal <CTD>									
Query Match 10.5%; Score 98; DB 2; Length 593;									
Best Local Similarity 25.3%; Pred. No. 1,8; Matches 48; Conservative 25; Mismatches 45; Indels 72; Gaps 8;									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 44 TPIIRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97									
Qy 10 TADVRNGIDGHALADRIGLDEIAWLSFGIDD----DTMALARAE-----53									
Db 41 TPIRRNGGE--LTNLIGNSSETTALARNEESATADLTAAVADTVAAAENAGAAWEA 97			</td						

C;Species: Halobacterium sp. NRC-1
 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C;Accession: A84110
 R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithäuser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablec Jung, K.H.; Alam, M.; Freitas, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A;Title: Genome sequence of Halobacterium species NRC-1.
 A;Reference number: A84160; MUID:20304483; PMID:11016950
 A;Accession: A84210
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-883 <STO>
 A;Cross-references: GB:AE004437; PIDN:910580117; PIDN:AGI19045.1; GSPDB:GN00138
 A;Gene: VNG0514C
 C;Genetics:
 Query Match 10.1%; Score 94.5; DB 2; Length 883;
 Best Local Similarity 28.3%; Pred. No. 5.8;
 Matches 51; Conservative 23; Mismatches 63; Indels 43; Gaps 10;
 QY 12 DVRNGIDGH - ALADRGILDDEAETANRLLSFTGIDDDTMALAAEQLFEA-----TAD 62
 Db 180 DVKSNTVEGQQLDLRQTA- -DKEAADPHDRDLASHNTALAEVTDIEHFEAEREQARQTRD 237
 QY 63 ALVTDYDFDHLSSEYERQDLFANSTKTEVQLEKETOAYLLGGRGEYDTEYAAQRA---- 117
 Db 238 ---DAADVLERYEESERTTALADEETTAIDREAVAE-----AEEPRETIAADRYSDHRE 286
 QY 118 RIGKTHDV-----LGL-GPDVYLGATRYTGTGILDAAD- -DVAADRGEEAAADEV 167
 Db 287 RASDLDDEAAALAAADIGLDDDAEDAASAE-----DAVAQREAEAERYREVAPAVSRU 340
 QY 441 PSYEBQTPATPDCFTSFVDLHLHRSSFNAPQQLZAIHWAMHTAGTSSGVKK-QBWPPTL 499
 QY 104 GRGEYDTEYYAAQRARIGKTHDVYLGLGDVYLGATRYTGTGILDAAD- -D 150
 Db 500 VQGPPGT-----GKTHHTWGMNVHLVQOQYTTSLLKKAPETYNQANECCSSD 550
 QY 151 DVAADRGEEAAADEV 175
 Db 551 NTLSGSIDEVQNMDNLFRTLPLK 575

RESULT 14
 C84609 hypothetical protein At2g22130 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Feb-2001 #text_change 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001 #sequence_revision 02-Feb-2001
 C;Accession: C84609
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fuji, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: C84609
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-2048 <STO>
 A;Cross-references: GB:AE002093; PIDN:96598689; PIDN:AGI18668.1; GSPDB:GN00139
 A;Gene: At2g22130
 A;Map position: 2

Query Match 10.0%; Score 93.5; DB 2; Length 2048;
 Best Local Similarity 26.3%; Pred. No. 20;
 Matches 55; Conservative 25; Mismatches 78; Indels 51; Gaps 11;

QY 3 NDNDTIVTADYR-NGID- -GHALADRIGL---DEAEIAMRLSFTGIDDDTMALAAEQP 55
 Db 1302 NPSRALAVADEMNAVDVLCRILSSNNYTMELKGDAAEELCYVLFANTIRIRSTVAACVEP 1361
 QY 56 LFEEATA DALVTDYDFDHLESVERTQDLFANSTKTEVQLEKETOAY---- -LLGLGRGE-YD 109
 Db 1362 L---VSLLVTEFSPAQHSVSYVRALDKLVEDD---EQLAEILVAARHGAVVPLVGLLYGVNV 1413
 QY 110 TEYAAQRA--RIGK-----IHDLVGLGPDPDVYLGAATRYTGTGILDAAD 150
 Db 1414 LHEAISRALVKLGDRPACKLEMVKAGVLDCLVFADLFCAAFSE---LLRILTN 1469
 QY 151 DVAADRGEEAAADEV 179
 Db 1470 NATIAKGQSAAKVE-----PLFHLLT 1491

RESULT 15
 T00533 probable DNA2-NAM7 helicase family protein [imported] - Arabidopsis thaliana
 N;Alternate names: SEN1 protein homolog T20K24.14
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 16-Feb-2001
 C;Accession: T00533; GB5752
 R;Rounsley, S.D.; Kaul, S.; Lin, X.; Cronin, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, July 1997
 A;Description: Arabidopsis thaliana chromosome II BAC T20K24 genomic sequence.
 A;Reference number: Z14167
 A;Accession: T00533
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1090 <ROU>
 A;Cross-references: EMBL:AC002392; PIDN:93176701; PID:93176714
 A;Experimental source: cultivar Columbia
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fuji, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A;Reference number: A84420; MUID:20083487; PMID:10617197
 A;Accession: G84572
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1090 <STO>
 A;Cross-references: GB:AB002093; PIDN:AGI12029.1; GSPDB:GN00139
 C;Genetics:
 A;Gene: T20K24.14; At2g19120
 A;Map position: 2
 A;Introns: 250/1; 310/1; 342/1; 383/2; 427/3; 1016/3
 Query Match 10.1%; Score 94; DB 2; Length 1090;
 Best Local Similarity 26.9%; Pred. No. 83;
 Matches 39; Conservative 17; Mismatches 55; Indels 34; Gaps 5;

QY 55 PLFE---ATADALVTDFYDHLSEYERTQDLFA-----NSIKTVQELKETQAEYLGL 103

AA; Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu0027

A:Map position: circular chromosome						
	Query Match	Score	DB	Length	881;	
	Best Local Similarity	10.0%	Score	93;	DB	2;
	Pred.	28.7%	No.	7.7;		
	Matches	49;	Mismatches	23;		
	Conservative		Indels	57;		
	Matches		Gaps	42;		
Qy	37 LSFTGIDDTMVALAAQPLFEAT---	10.0%	ADALVTDYFDHLSSEYERTDIL---	10.0%	-PANST	86
Db	621 LTFVNNTD---SVRAERALKENDAKDELKNDFQHV-SYELRSPLTNITGFTDIL					675
Qy	87 KT- -VIEQLKETOAQYLGLGRGYDTEAAQRARIKGHDVLGIPGDYLG-AYTRYTGT	10.0%				143
Db	676 RTPGIGGOLTERAAYL-----DHISTSSSVLTINDLLA-TVDAGINQLYNSDN					726
Qy	144 LDLDALADV--VADRGEAAAAVDVELVARYLP-	10.0%	MLKLKT	179		
Db	727 DLNEFLDDSVSQTADRLQESGTSILEVPAHLGSSLVAHQRLKQILRKT					777

Search completed: August 10, 2004, 15:29:28
Job time : 9.01937 secs

Result No.	Score	Query Match	Length	DB ID	Description
1	933	100.0	489	1 HMAT_HALN1	Q9hp6 halobacteri
2	933	100.0	489	1 HMAT_HALSA	P71433 halobacteri
3	142	15.2	432	1 HMAT_BACSU	P07621 bacillus su
4	99	10.6	448	1 SPG1_STRSG	P06634 streptococc
5	98	10.5	593	1 SPG2_STRSG	P19939 streptococc
6	90.5	9.5	305	1 DPX_RALSO	Q9kh16 ralstonia s
7	89	9.5	955	1 KINL_LEICH	P46850 mycobacteri
8	88	9.4	775	1 PLB5_MYCLE	Q9x700 halobacteri
9	87	9.3	469	1 TTG_MYCLE	Q9cbx2 mycobacteri
10	85.5	9.2	957	1 SECA_MYCSM	P71533 mycobacteri
11	82.5	8.8	381	1 GVP_C_HALME	P02238 halobacteri
12	82	8.8	331	1 VAI_BPSP	P09677 bacterioph
13	82	8.8	372	1 PROB_PSEAE	Q9hv19 pseudomonas
14	81	8.7	1001	1 IF2_SYN3	P72669 synechocyst
15	80.5	8.6	764	1 HTR2_HALSA	P71410 halobacteri
16	80.5	8.6	765	1 DPP4_BOVIN	P81425 bos taurus
17	80.5	8.6	765	1 DPP4_TELCA	Q9n217 felis silv
18	80.5	8.6	867	1 SYA_FUSNN	Q8rxj8 fusobacteri
19	80	8.6	461	1 K6PF_THEZI	Q9nhb1 thermococcu
20	80	8.6	954	1 GCSP_SSISM	Q88715 pseudomonas
21	79.5	8.5	334	1 YJH3_RHIME	P92v44 rhizobium m
22	79.5	8.5	763	1 HTR2_HALN1	Q9np81 halobacteri
23	79.5	8.5	1415	1 RPOC_HAEIN	P43739 haemophilus
24	78.5	8.4	374	1 PROB_RALSO	P8rxv1 halostonia s
25	78.5	8.4	382	1 GVCL_HALN1	P24574 halobacteri
26	78.5	8.4	454	1 K6PF_DYRFU	Q9vz7 pyrococcus
27	78.5	8.4	584	1 Y742_STRCO	P2748 homo sapien
28	78.5	8.4	766	1 DPP4_HUMAN	Q05858 mycobacteri
29	78.5	8.4	949	1 SEAL_NYCTU	P80153 xanthomonas
30	78	8.4	442	1 HRB6_XANCV	Q59315 pyrococcus
31	78	8.4	450	1 K6PF_PYRHO	Q97793 thermococcu
32	77.5	8.3	459	1 K6PF_THELI	P70985 bacillus st
33	77	8.3	323	1 T2B1_BACST	

Scoring table: BLOSUM62
Searched: Gapop 10.0 , Gapext 0.5
Total number of hits satisfying chosen parameters: 141681
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_42;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Total number of hits satisfying chosen parameters: 141681
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

ALIGNMENTS

RESULT 1
HMAT_HALN1
ID HMAT_HALN1
AC Q9HP66;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Heme-based aerotactic transducer hemAT.
GN HEMAT OR HTR10 OR VNG150SG
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE-20504433; PubMed-11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrigna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leitbauer B., Keller K., Cruz R., Danion M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krabs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebbhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC |- FUNCTION: Heme-containing signal transducer responsible for aerotaxis the migration response toward or away from oxygen (By similarity).
CC |- SUBUNIT: Homotrimer (Probable).
CC |- SIMILARITY: Contains 1 methyl-accepting transducer domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC DR EMBL; AE005064; AUG19801.1; -.
CC DR HSSP; EB3304; E84304.
CC DR HSSP; P02942; 1QU7.
CC DR InterPro; IPR04089; ChmotaXis transd.
CC DR InterPro; IPK004030; Me chmotaXis.
CC DR PRINTS; PR000151; MCPsignat_1.
CC DR SMART; SM0283; MA_1.
CC DR PROSITE; PSS0111; CHEMOTAXIS_TRANSDUC_2; 1.
CC KW Transducer; Hemme; Complete proteome.
CC FT DOMAIN 218 454 METHYL_ACCEPTING TRANSDUCER.
CC SQ SEQUENCE 489 AA; 52852 MW; 8COAF17966791E CR64;
CC Query Match 100.0%; Score 933; DB 1; Length 489;
CC Best Local Similarity 100.0%; Pred. No. 4.1e-68;
CC Matches 184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSNDNDTLLTADVRNGIDGHALADRIGLDAEIAWRLSFTGIDDITMAAIAEQPLFEAT	KW	Transducer; Heme.
Db	1	MSNDNDTLLTADVRNGIDGHALADRIGLDAEIAWRLSFTGIDDITMAAIAEQPLFEAT	ET	218 454 METHYL-ACCEPTING TRANSDUCER.
DB	61	ADALVTDYDILESTERTQLDFANSTKTVQELKEAQEYLGLGRGEYDTEYAQRARIG	FT	SEQUENCE 499 AA; 52816 MW;
Qy	61	ADALVTDYDILESTERTQLDFANSTKTVQELKEAQEYLGLGRGEYDTEYAQRARIG	SQ	Score 933; DB 1; Length 489;
Db	61	ADALVTDYDILESTERTQLDFANSTKTVQELKEAQEYLGLGRGEYDTEYAQRARIG	Query Match	100.0%; Score 100.0%; Pre 4.1e-68;
Db	121	KIHDVTLGLGPDVYLGAATRYTYTGGLDALADDVVADELRGEAAAVDELVARFLPMLKLTF	Best Local Similarity	100.0%; Pre 0; Mismatches 0;
Qy	121	KIHDVTLGLGPDVYLGAATRYTYTGGLDALADDVVADELRGEAAAVDELVARFLPMLKLTF	Matches	Gaps 0;
Db	121	KIHDVTLGLGPDVYLGAATRYTYTGGLDALADDVVADELRGEAAAVDELVARFLPMLKLTF	Qy	1 MSNDNDTLLTADVRNGIDGHALADRIGLDAEIAWRLSFTGIDDITMAAIAEQPLFEAT 60
Qy	181	DQQI 184	Db	1 MSNDNDTLLTADVRNGIDGHALADRIGLDAEIAWRLSFTGIDDITMAAIAEQPLFEAT 60
Db	181	DQQI 184	Qy	61 ADALVTDYDILESTERTQLDFANSTKTVQELKEAQEYLGLGRGEYDTEYAQRARIG 120
RESUL T 2			Db	61 ADALVTDYDILESTERTQLDFANSTKTVQELKEAQEYLGLGRGEYDTEYAQRARIG 120
HMAT_HALSA			Qy	61 ADALVTDYDILESTERTQLDFANSTKTVQELKEAQEYLGLGRGEYDTEYAQRARIG 120
ID			Db	121 KIHDVTLGLGPDVYLGAATRYTYTGGLDALADDVVADELRGEAAAVDELVARFLPMLKLTF 180
HMAT_HALSA		STANDARD;	Qy	121 KIHDVTLGLGPDVYLGAATRYTYTGGLDALADDVVADELRGEAAAVDELVARFLPMLKLTF 180
AC	P71413;	PRT;	Db	121 KIHDVTLGLGPDVYLGAATRYTYTGGLDALADDVVADELRGEAAAVDELVARFLPMLKLTF 180
DT	28-FEB-2003	(Rel. 41, Created)	Qy	181 DQQI 184
DT	28-FEB-2003	(Rel. 41, Last sequence update)	Db	181 DQQI 184
DT	28-FEB-2003	(Rel. 41, Last annotation update)	Qy	181 DQQI 184
DE	Heme-based aerotactic transducer hemAT (Transducer HtB protein).			
GN	HEMAT OR HtB.		DB	181 DQQI 184
OS	Halobacterium salinarum.		Qy	181 DQQI 184
OC	Euryarchaeota; Halobacteria; Halobacteriales;		Db	181 DQQI 184
OC	Halobacteriaceae; Halobacterium.		Qy	181 DQQI 184
OX	NCBI_TaxID:2242;		Db	181 DQQI 184
RN	[1]		Qy	181 DQQI 184
RP	SEQUENCE FROM N.A.		Db	181 DQQI 184
RC	STRAIN=FLX15;		Qy	181 DQQI 184
RX	MEDLINE=96209786; PubMed=8643458;		Db	181 DQQI 184
RA	Zhang W., Brooun A., McCandless J., Banda P., Alam M.,		Qy	181 DQQI 184
RA	"Signal transduction in the archaeon Halobacterium salinarum is processed through three subfamilies of 13 soluble and membrane-bound transduc		Db	181 DQQI 184
RA	transducer proteins."		Qy	181 DQQI 184
RA	Proc. Natl. Acad. Sci. U.S.A. 93:4649-4654 (1996).		Db	181 DQQI 184
RN	[2]		Qy	181 DQQI 184
RP	FUNCTION.		Db	181 DQQI 184
RX	MEDLINE=20140131; PubMed=10676961;		Qy	181 DQQI 184
RA	Hou S., Larsen R. W., Boudko D., Riley C. W., Karatan E., Zimmer M., Ordal G. W., Alam M.,		Db	181 DQQI 184
RA	"Myoglobin-like aerotaxis transducers in Archaea and Bacteria.";		Qy	181 DQQI 184
RA	Nature 403:540-544 (2000).		Db	181 DQQI 184
RN	-1- FUNCTION: Heme-containing signal transducer responsible for aerotaxis, the migratory response toward or away from oxygen.		Qy	181 DQQI 184
CC	-1- SUBUNIT: Homotrimer (Probable).		Db	181 DQQI 184
CC	-1- SIMILARITY: Contains 1 methyl-accepting transducer domain.		Qy	181 DQQI 184
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).		RA	98044033; PubMed=9384377;
CC	-----		RA	MEDLINE=98240224; PubMed=9579061;
CC	-----		RA	NCBI_TaxID:98; PMID:15200000;
CC	-----		RA	NCBI_TaxID:1423;
CC	-----		RA	SEQUENCE FROM N.A.
CC	-----		RA	STRAIN=168;
CC	-----		RA	RX
CC	-----		RA	NCBI_TaxID:98; PMID:15200000;
CC	-----		RA	NCBI_TaxID:1423;
CC	-----		RA	SEQUENCE FROM N.A.
CC	-----		RA	RX
CC	-----		RA	MEDLINE=98044033; PubMed=9384377;
CC	-----		RA	Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G., Azvedo V., Boller T.M.G., Bessieres P., Bolotin A., Borchart S., Brignell S.C., Bron S., Borriss R., Bourquier L., Brans A., Braun M., Bruneau S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guisseppi G., Guy B.J., Haga K., Haeich J., Harwood C.R., Hillaire B., Holzapfel S., Hosono S., Hullo M.F., Itaya M., Klein C., Kojris B., Karamata D., Kashiwa S., Klaerr-Bianchard M., Krogsgaard M., Kuroda T., Koertner P., Koningsstein G., Krogh S., Kumano M., Kobayashi Y., Koertner P., Koningsstein G., Krogsgaard M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maeul C., Medina N., Mellado R. P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portwaille B., Rapoport G., Rey M., Reynolds S., Presecan E., Pujić P., Purcell B., Rivoal C., Rocha E., Rieger M., Rivolta C., Schröter R., Sadaie Y., Sato T., Scanlan B., Schleich S., Scoffone F.,

RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.-S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi K., Takei M., Takei K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terestra F., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandembol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambut R., Wedler B., Weitznerger T.,
 RA Winters P., Wipat A., Yamamoto H., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshioka H.F., Zumstein E., Yoshiwara H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*".
 RT [3]
 RL Nature 390:249-256 (1997).
 RN
 FUNCTION.
 RP
 RX MEDLINE=20140131; PubMed=10676961;
 RA Hou S., Larsen R.W., Boudko D., Riley C.W., Karatan E., Zimmer M.,
 RA Ordal G.W., Alam M.;
 RA "Myoglobin-like aerotaxis transducers in Archaea and Bacteria.";
 RL Nature 403:340-344 (2000).
 RN
 J. [4]
 RN
 RESONANCE SPECTROSCOPY, AND SUBUNIT.
 RX MEDLINE=21950695; PubMed=11821422;
 RA Aono S., Kato T., Matsuki M., Nakajima H., Ohta T., Uchida T.,
 RA Kitagawa T.;
 RT "Resonance Raman and ligand binding studies of the oxygen-sensing
 signal transducer protein HemAT from *Bacillus subtilis*".
 RT J. Biol. Chem. 277:1528-13538 (2002).
 CC -!- FUNCTION: Heme-containing signal transducer responsible for
 aerotaxis, the migratory response toward or away from oxygen.
 CC -!- SIMILARITY: Contains 1 methyl-accepting transducer domain.
 CC -!- This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).
 CC or send an email to license@isb-sib.ch).

EMBL: Y14084; CAA7445.1.;
 DR InterPro: IPR004089; CAB12878.1.;
 DR PIR: C69832; C69832. [5]
 DR SubList: BG13066; hemat.
 DR GO: GO:0004871; F: heme binding; IDA.
 DR GO: GO:009454; P:aerotaxis; IDA.
 DR GO: GO:00015; MCP:signal; 1.
 DR SMART: SM00283; MA: 1.
 DR PROSITE: PS50111; CHEMOPAXIS TRANSDUC_2; 1.
 KW Transducer; Heme; Complete proteome.
 FT DOMAIN 184 42 AA; 48767 MW; 87D0655374070B64 CRC64;
 SQ SEQUENCE 432 AA; Score 142; DB 1; Length 432;
 Best Local Similarity 25.3%; Score 0.00027; Mismatches 76; Indels 10; Gaps 2;
 Matches 39; Conservative 29; Mismatches 76; Indels 10; Gaps 2;

Query Match 31 AEIAWRLSPGIDDMMALAAEQLPFEATADALYTVDFYDHLSEYERTQDLPANSTKTYE
 QY :::::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:
 Db 32 ADVKKQLKAVTRLGDAELYVQEQLIQQENVNIVDAFYRNLDHESSLMDDII-NDHSSYD 90
 QY 91 QLKETQAELVGLGRGEYDTTEYAAQRARIGKIHDVYLGLGPDVYLGAYTRYTGILDALAD 150
 Db 91 RLKQTLKRHIIQEMPAGVIDDEFIEKRNRASITHRIGLPKWVNGAFQELLLSMIDY-- 148
 QY 151 DVVADRGEZAAADEVDELVAFPLMLKLTDQOQI 184
 Db 149 -----EASITNQELLKXIKATTKLNLEQQQ 175

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M13825; AAA03664.1.;
 DR PIR; A24496; A24496.
 DR PDB; 1EN7; 08-MAY-02.
 DR PDB; 1F0C; 20-JUL-95.
 DR PDB; 1G81; 15-APR-93.
 DR PDB; 1IGC; 03-JUN-95.
 DR PDB; 1IGD; 01-NOV-94.
 DR PDB; 1IME; 30-OCT-02.

DR	PDB; 1MVK; 30-OCT-02.	ID	SPG2 STRSG	STANDARD;	PRT;	593 AA.
DR	PDB; 1PGA; 30-APR-94.	AC	P19009;			
DR	PDB; 1PGB; 30-APR-94.	DT	01-FEB-1991 (Rel. 17, Created)			
DR	PDB; 1PGX; 15-JUL-92.	DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DR	PDB; 2GB1; 15-APR-93.	DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DR	PDB; 2IGD; 29-JUL-98.	DE	Immunoglobulin G binding protein G precursor (Igg binding protein G).			
DR	PDB; 2IGH; 31-JAN-94.	GN	SPG.			
DR	InterPro; IPR002988; GA.	OS	Streptococcus sp. (Lancefield group G).			
DR	InterPro; IPR002988; Gram_pos_anchor.	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
DR	InterPro; IPR001899; Gram_pos_anchor.	OC	Streptococcus.			
DR	InterPro; IPR000724; IgG_Bind_B.	OC	NCBI_TaxID=1320;			
DR	Pfam; PF01468; GA; 2.	RN	[1]_TaxID=1320;			
DR	Pfam; PF00746; Gram_pos_anchor; 1.	RP	SEQUENCE FROM N.A.			
DR	Pfam; PF01378; IgG_Binding_B; 2.	RC	STRAIN=G148;			
DR	Pfam; PF04650; YSIRK signal; 1.	RX	MEDLINE=88029445; PubMed=3665928;			
DR	PRINTS; PRO0015; GPOSANCHOR.	RA	Olsson A., Eliasson M., Guss B., Nilsson B., Hellman U.,			
DR	TIGRFAMS; TIGR0117; LPXTG_anchor; 1.	RA	Lindberg M., Uhlen M.;			
DR	TIGRFAMS; TIGR01168; YSIRK_anchor; 1.	RT	"Structure and evolution of the repetitive gene encoding streptococcal protein G";			
DR	PROSITE; PS50847; GRAM_POS_ANCHORING; 1.	RT	Eur. J. Biochem. 168:319-324 (1987).			
DR	IGG-binding protein; Cell_wall; Peptidoglycan-anchor; Repeat; Signal;	RL	[2]			
KW	3D-structure.	RN				
FT	SIGNAL; 1	RN				
FT	CHAIN; 34	33	IMMUNOGLOBULIN G BINDING PROTEIN G.			
FT	PROPEP; 418	417	REMOVED BY SORTASE (POTENTIAL).			
FT	DOMAIN; 69	448	ALA-RICH.			
FT	DOMAIN; 104	111	"Nucleotide sequence of the protein G gene from Streptococcus GX7805, and comparison to previously reported sequences."			
FT	REPEAT; 104	215	RT			
FT	REPEAT; 179	215	RT			
FT	DOMAIN; 228	352	RL			
FT	REPEAT; 228	352	RN			
FT	REPEAT; 298	352	SEQUENCE OF 114-593 FROM N.A.			
FT	DOMAIN; 386	410	RC			
FT	SITE; 414	5 X 5 AA REPEATS OF [DE]-D-A-K-K.	RX			
FT	MOD_RES; 417	418	RX			
FT	STRAND; 228	417	Guss B., Eliasson M., Olson A., Frej A.-K., Joernvall H., Flock J.-I., Lindberg M.;			
FT	STRAND; 239	234	RT			
FT	HELIX; 249	262	RT			
FT	TURN; 263	264	RL			
FT	STRAND; 268	272	RN			
FT	TURN; 273	276	STRUCTURE BY NMR OF 371-427.			
FT	STRAND; 277	281	RC			
FT	STRAND; 297	304	RX			
FT	STRAND; 309	316	MEDLINE=98290449; PubMed=6228455;			
FT	HELIX; 319	332	RA			
FT	TURN; 333	334	Malakauskas S.M., Mayo S.L.;			
FT	STRAND; 338	342	Design, structure and stability of a hyperthermophilic protein			
FT	TURN; 343	346	RT			
FT	STRAND; 347	351	Variant."			
SEQ	SEQUENCE; 448 AA;	47567 NW;	RL			
QY	10 TADVRNGTDGHALADRIGLDEAETAWRSRFTGIDD-----DTMAAALAAE-----	53	Nat. Struct. Biol. 5:470-475 (1998).			
DB	41 TPIRNGT-----LTMILGNSSETLAIIRNEESATADLTAAVADTVAAENAGAAAEAA 97	CC	-!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by an amide bond (Potential).			
QY	54 ----QPLFEEATDAL-----VTFDFYDHELEYERTQLFLANSRFTVTLKEYTQAYLIG 102	CC	CC			
DB	98 AAAADALAKAKADALKEFNKYGSDYKKNL-----INNAKTVGIKDLQAQVV - 145	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the BMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
QY	103 LGRCGYDTYYAQPDRIGKHDVVLGLPDPYLGAYTRYTGTGLDALADDVADRGEEAAA 162	CC	CC			
DB	146 -----ESAKKARISSEATD-----GLSDFLKSQTTPA--EDTVK 175	CC	CC			
QY	163 AVDELVARFL 172	CC	CC			
DB	176 SIEELAAEKVL 185	CC	CC			
RESULT	SPG2_STRSG	CC	CC			

DR	Pfam; PF04650; YSIRK_SIGNAL; 1.	SEQUENCE FROM N.A.	
DR	PRINTFAMS; TIGR01167; LPXTG_anchor; 1.	RC STRAIN=GM11000;	
DR	TIGR01168; YSIRK_SIGNAL; 1.	RC MEDLINE=21681879; PubMed=11823852;	
DR	GRAM_POS_ANCHORING; 1.	RX RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,	
DR	IGG-binding protein; Cell wall; Peptidoglycan-anchor; Repeat; Signal;	RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,	
KW	3D-structure.	RA Chandler M., Choisne N., Claude R., Demange N.,	
KW	FT SIGNAL 1 33	RA Gaspin C., Lavie M., Moian A., Robert C., Saurin W., Schiex T.,	
FT	CHAIN 34 562 IMMUNOGLOBULIN_G_BINDING_PROTEIN_G.	RA Signer P., Thebaud P., Whalen M., Wincker P., Levy M.,	
FT	PROPEP 563 593 REMOVED_BY_SORTASE_(POTENTIAL).	RA Weissbach J., Boucher C.A.; RT "Genome sequence of the plant pathogen <i>Ralstonia solanacearum</i> ."	
FT	DOMAIN 69 111 ALA-RICH.	RL Nature 415:497-502(2002)	
FT	DOMAIN 104 290 3 X 37 AA REPEATS.	-!- FUNCTION: Involved in the biosynthesis of Lipid A, a	
FT	REPEAT 104 140 1-1.	CC phosphorylated glycolipid that anchors the lipopolysaccharide to	
FT	REPEAT 179 215 1-2.	CC the outer membrane of the cell (By similarity).	
FT	REPEAT 254 290 1-3.	CC -!- CATALYTIC ACTIVITY: UDP-3-O-(3-hydroxytetradecanoyl)-N-	
FT	DOMAIN 303 427 2 X 55 AA REPEATS.	CC acetylglucosamine + H(2)O = UDP-3-O-(3-hydroxytetradecanoyl)-N-	
FT	REPEAT 303 357 2-2.	CC Glucosamine + acetate.	
FT	REPEAT 373 427 2-2.	-!- SIMILARITY: Belongs to the lpxC family.	
FT	DOMAIN 531 555 5 X 5 AA REPEATS OF [DE]-D-A-K-K.	CC -!- PATHWAY: Lipid A biosynthesis; second step.	
FT	SITE 559 563 LPXTG SORTING SIGNAL (POTENTIAL).	CC	
FT	MOD RES 562 562 AMIDE-LINKED TO CELL WALL (POTENTIAL).	CC	
FT	STRAND 373 379	CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
FT	STRAND 386 386	CC between the Swiss Institute of Bioinformatics and the EMBL outstation -	
FT	STRAND 389 390	CC the European Bioinformatics Institute.	
FT	HELIX 394 408	CC There are no restrictions on its	
FT	TURN 409 409	CC use by non-profit institutions as long as its content is in no way	
FT	STRAND 417 417	CC modified and this statement is not removed. Usage by and for commercial	
FT	TURN 418 421	CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
FT	STRAND 422 426	CC or send an email to license@isb-sib.ch).	
SQ	SEQUENCE 593 AA; 63291 MW; 048BA760D5B2920 CRC64;	CC	
DR	EMBL; AF254622; AACF7524_1; -.	CC	
DR	EMBL; AL646072; CAD16544_1; -.	CC	
DR	HAMAP; MF_00388; -; 1.	CC	
PFam	PF03331; LpxC; 1.	CC	
DR	TIGR0325; lpxC; 1.	CC	
KW	Hydrolase; Lipid_A_biosynthesis; Lipid synthesis; Complete proteome.	CC	
FT	CONFFLICT 35 N -> T (IN REF. 1).	CC	
FT	CONFFLICT 132 132 I -> T (IN REF. 1).	CC	
FT	CONFFLICT 155 155 S -> A (IN REF. 1).	CC	
FT	CONFFLICT 279 279 A -> T (IN REF. 1).	CC	
SQ	SEQUENCE 305 AA; F531BBE1349FBAB CRC64;	CC	
Query Match	9.7%; Score 90.5; DB 1; Length 305;	Query Match	9.7%; Score 90.5; DB 1; Length 305;
Best Local Similarity	25.3%; Pred. No. 1.4; Gaps 8;	Best Local Similarity	24.6%; Pred. No. 2.6; Gaps 10;
Matches	48; Conservative 25; Mismatches 45; Indels 72; Gaps 8;	Matches	50; Conservative 20; Mismatches 58; Indels 75; Gaps 10;
Qy	10 TADVRNGIDGHALADRIGLDFAEIARWLSTFGIDD-----DTMAALAAE----- 53	Qy	12 DVRNGIDGHALADRIGLDFAEIARWLSTFGIDD TMAALAAQPLFEATADALVTDFYDH 71
Db	41 TPIIRNGGE---LTNLJGNSETTLARNEESATADLTAAAVADTVAAAENAGAAWEA 97	Db	136 EVRDG-----DKLARLEPYFGFKLSFT-----IEFRHADVRTGQTFFEDFADT 179
Qy	54 ---QPLFEATADAL-----VTDYDFHIESYERTQDLANSTKTEQKETQAEYL LG 102	Qy	72 LESYERTQDLFANSKTQKETQAEYL LG 102
Db	98 AAAADALAKAKADAKEFNKYGVSDYKNL-----INNAKTYEVGKDLOQQVV- 145	Db	180 SYTRBIAARTPFGFHAEVEMLR-----GLARGSSLDNAIVDELRYGDE 233
Qy	103 LGRGYDTEAQQRARIKGKHDVLGLGDPYLGATRYTTGLLDAADDVVAADREAAA 162	Qy	112 YAAQRARIKGKHDVLGLGDPYV----LGATRYTTGLLDAADDVVAADREAAA 162
Db	146 -----ESAKKARISEATD-----GLSDFELIKSQTTPA--EDTVK 175	Db	234 FVRH----KILDATG---DLYVYGPPLAAYTAHKSGHCLNNALLRAJADEAYEVYT 285
Qy	163 AVDELVARFL 172	Qy	154 ADRGEEAAAVADVELVARFLPMLK 176
Db	176 SIELAAKVL 185	Db	286 FDKVVEAPRA-----FLPQLQ 301
RESULT 6	Bacteria; Ralstonia (Pseudomonas solanacearum); Burkholderiales; Burkholderiaceae; Ralstonia.	RESULT 7	KIN1_leigh STANDARD; PRT; 955 AA.
LPXC_RALSO	ID KIN1_leigh	Qy	AC P4665; 1-32, Created
ID LPXC_RALSO	STANDARD;	Db	AC 01-NOV-1995 (Rel. 32, Created)
AC Q9KH16;	PRT; 305 AA.	DT 01-OCT-1995 (Rel. 32, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Created)	RN SEQUENCE FROM N.A.	DT 01-OCT-1996 (Rel. 34, Last annotation update)	
DT 28-FEB-2003 (Rel. 41, Last sequence update)	STRAIN=ACH0158 / Bt-ovar 2;	DE Kinesin-like protein K39 (Fragment).	
DE UDP-3-O-(3-hydroxymyristoyl) N-acetylglucosamine deacetylase	RA "Putative lpxC gene from <i>Ralstonia solanacearum</i> ACH0158 strain."	GN OS Leishmania chagasi.	
DE (EC 3.5.1.-) (UDP-3-O-acyl-GlcNAc deacetylase)	RA Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.	NCBI_TaxID=305;	
GN LPXC_OR_RSC2837 OR RS00268.	RA [2]	OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.	
OS Ralstonia solanacearum (Pseudomonas solanacearum)	RA [1]	OC	
OC Bacteria; Betaproteobacteria; Burkholderiales;	RA [1]	OC	
OC Burkholderiaceae; Ralstonia.	RA [1]	OC	
OX NCBI_TaxID=305;	RA [1]	OC	
RN SEQUENCE FROM N.A.	RA [1]	OC	
RP STRAIN=ACH0158	RA [1]	OC	
RC "Putative lpxC gene from <i>Ralstonia solanacearum</i> ACH0158 strain."	RA [1]	OC	
RL Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.	RA [2]	OC	

OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.	AC	P71533;
OX	[1] - TaxID=1769;	DT	15-JUL-1998 (Rel. 36, Created)
RN	SEQUENCE FROM N.A.	DT	15-JUL-1998 (Rel. 36, Last sequence update)
RP	STRAIN=IN;	DE	30-MAY-2000 (Rel. 39, Last annotation update)
RC	MEDLINE=2111228732; PubMed=11234002;	GN	Preprotein translocase secA subunit.
RX	Cole S.T., Eigmeyer K., Parkhill J., James K.D., Thomson N.R., Wheeler N., Honore N., Garnier T., Churcher C., Harris D., Mungall K., Basham D., Brown D., Chillingworth T., Connor P., Davies R.M., Devlin K., Duthoy S., Fellwell T., Fraser A., Hamlin N., Hollroyd S., Hornsby T., Jagele K., Lacour M.A., Maclean J., Moule S., Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M., Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Taylor K., Whittlehead S., Woodward J.R., Barrell B.G.;	OS	Mycobacterium smegmatis.
RA	"Massive gene decay in the leprosy bacillus.";	OC	Bacteria; Actinobacteria; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
RA	Nature 403:1007-1011(2000).	OC	Myco bacterium smegmatis .
CC	-!- FUNCTION: Involved in protein export. Acts as a chaperone by maintaining the newly synthesized protein in an open conformation	OX	Bacterium; Actinobacteridae; Mycobacterium.
CC	-!- SIMILARITY: Belongs to the FKBP-type PPIase family. Tig subfamily.	RN	SEQUENCE FROM N.A.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	RP	Brown A.B., Jacobs W.R.;
CC	EMBL; AL083922; CAC30431.1; -.	RA	"Conservation of the general secretory pathway: isolation and characterization of SecA homologues from Mycobacterium bovis BCG and M. smegmatis";
CC	RL	RT	Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
CC	-!- FUNCTION: Involved in protein export. Interacts with the secY/secE subunits. SecA has a central role in coupling the hydrolysis of ATP to the transfer of pre-secretory periplasmic and outer membrane proteins across the membrane (BY similarity).	CC	-!- SUBUNIT: Part of the prokaryotic protein translocation apparatus which comprise secA, secB, secC, secD, secE, secF, secG and secY (BY similarity).
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic side of membrane (BY similarity).	CC	-!- SIMILARITY: Belongs to the secA family.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	PIR; B87094; B87094.	DR	EMBL; U66081; AAB0754.1; -.
DR	Leproma; MF_00303; -_1.	DR	InterPro; IPR001650; SecA.
DR	HAVAP; MF_00303; -_1.	DR	InterPro; IPR00165; SecA.
DR	InterPro; IPR001179; FKBP_PPIase.	DR	Pfam; PR00271; helicase C. 1.
DR	InterPro; IPR005215; Tig_fac.	DR	Pfam; PF01043; SecA_protein; 1.
DR	InterPro; IPR008880; Trigger_C.	DR	PRINTS; PR00906; SEC_A.
DR	InterPro; IPR008881; Trigger_N.	DR	TIGRFAMS; TIGR00963; secA; 1.
DR	Pfam; PF00254; FKBP; 1.	DR	PROSITE; PS01312; SECA.
DR	Pfam; PF05698; Trigger_C; 1.	FT	Protein transport; ATP binding; Membrane; Transport.
DR	Pfam; PF05697; Trigger_N; 1.	NP BIND	NP BIND 98 AA; 105 ATP (POTENTIAL).
DR	TIGRFAMS; TIGR00115; Tig; 1.	SQ	SEQUENCE 957 AA; 107001 MW; AD208569A22BA32E CRC64;
DR	PROSITE; PS00453; FKBP_PPIASE_1; FALSE_NEG.	Query Match	9.2%; Score 85.5; DB 1; Length 957;
DR	PROSITE; PS00454; FKBP_PPIASE_2; FALSE_NEG.	Best Local Similarity	21.8%; Pred. No. 24;
DR	PROSITE; PS00509; FKBP_PPIASE_3; 1; FALSE_NEG.	Matches	46; Conservative 32; Mismatches 70; Indels 63; Gaps 9;
DR	KW DOMAIN 162 243 Complete proteome.	Qy	4 DNDFLTVTADVRN-----GIDGHALAD-----RIGL-----DEAEIA 34
FT	SEQUENCE 469 AA; 51476 MW; PPIASE; FKBP_TYPE.	Db	350 ENQTLATITLQNYFRYDYLKLSGMGTAAEELHTRLYRGLGVPPINTKPMVYQDQSDLI 409
SQ	54095F5BD231B884 CRC64;	Qy	35 WR---LSFGTIDDDTMANAAQPLFEATADALVTDYFHDLESYERTODLFANSTKVTEQ 91
Qy	9 VTAIVRNRGIDGHAL---ADIGLDAEIAWRLSFTGIDDPTMAAEEQPLFEATADALVT 66	Db	410 YKTEBEAKFLAVVDVAERAKHAKGPVLIGTT-----SVERSEYSLKMLTKRVP 457
Db	165 VS1DLSATINGEKVPAWNADEAGLSH-EVGYGRLLAGLDLAVLGSAGESRF--TTQIAT 220	Qy	92 LKETQAEYLGLGRGEYDTEYAAQQRARIKTI---HDYIQLGLGPDVYLGAYTRYTGLLDAL 148
Qy	67 DFYDHLESYERTOOLFANSTKTVQLE-----TQAELYLLGGRGEY 108	Db	458 HNVLNK-----HEQANITAEGRGAVYATWNGRGTDIVG----GIVDFI 504
Db	221 -----SKHAGQD---AEVIVTVKSVKRELPDPDEFAQLVSEFDWMAELRNLGDQVR 271	Qy	149 ADDVYADRGEEAAAVDELVARE----LPLMLK 176
Qy	109 DTEYAAQRARI-----GKTHDVL-GLGDPDVYLGAYTRYY 141	Db	505 ADKRLRLRGLDPVETPEYEAAWHEVLDQVK 535
Db	272 KAKYAAQQAELKIRDAAVDALELERYDVLVPLPIGIVQQFNNAHLDAISGLHD-----321	RESULT 11	
Qy	142 TGLIDLAADDVADRG-----EAAAADVDELVAREFAERSAERDVTRL-LLDVVAADDKI 366	GIVPC_HALME	
Db	322 -----EAKFAVLAERGSSREEFEAARSAAERDVTRL-LLDVVAADDKI 366	ID_GIVPC_HALME	STANDARD;
		AC_Q0228;	PRT; 381 AA.

DT	15-MAR-2004 (Rel. 43, Last annotation update)
DT	01-JUL-1993 (Rel. 26, Created)
DT	01-JUL-1993 (Rel. 26, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DN	Gas vesicle protein C.
GN	Gvpc.
EN	Halobacteriaceae; Halobacteria; Halobacteriales;
OC	Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
NCBI_TAXID	2252;
RN	SEQUENCE FROM N.A.
RP	"Three different but related gene clusters encoding gas vesicles in halophilic archaea.";
RL	J. Mol. Biol. 227:586-592 (1992).
CC	-!- FUNCTION: May confer stability to the gas vesicle membranes. Gas vesicles are small, hollow, gas filled protein structures that are found in several microbial planktonic microorganisms. They allow the positioning of the organism at the favorable depth for gas vesicle membrane.
CC	-!- SUBCELLULAR LOCATION: Binds to the external surface of the gas vesicle membrane.
CC	-!- SIMILARITY: VARY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF CYANOBACTERIA.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; X64701; CAA05944.1; -.
CC	PIR; S28115; S28115.
CC	InterPro; IPR008639; Halo_GVPC.
CC	Pfam; PF05465; Halo_GVPC;_1.
CC	Gas vesicle; Repeat_KW
CC	Domin; 22 274 7 X APPROXIMATE TANDEM REPEATS.
CC	REPEAT 22 59 1.
CC	FT REPEAT 60 84 2.
CC	FT REPEAT 85 122 3.
CC	FT REPEAT 123 160 4.
CC	FT REPEAT 161 192 5.
CC	FT REPEAT 193 232 6.
CC	FT REPEAT 233 274 7.
CC	FT DOMAIN 246 335 42653 MW; 9FB48199D0303921 CRC64;
CC	SEQUENCE 381 AA; 42653 MW; 9FB48199D0303921 CRC64;
CC	Query Match Score 8.8%; Best Local Similarity 25.9%; Matches 45; Conservative 27; Mismatches 73; Indels 29; Gaps 9; Result 13
CC	DR PROB_PSEAE STANDARD; PRT; 372 AA.
CC	QY 1 ADVRNGIDGHALADRIGLDEAETIARNLSTGICDDDTMAAAEQLFEATA DALYTFYD 70 AC QHVLVS; (Rel. 40, Created)
CC	DB 81 ADRRDAFD--AYADIFATDVAAEMQDSLIAADDLRAEMDETHEAFDADFTD-- 135 DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC	QY 71 HLESVERTQDLFANSKTKTBOLKETQAEYLIGR-GIYDTEYAAQRATGKHDVLGLG 129 DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC	DB 136 -VATLRDVSDL----LTAISLQSEFVQSGEYNGYASEFGAD--1DQEHAVAAK 184 DE Glutamate 5-kinase (EC 2.7.2.11) (gamma-glutamyl kinase) (GK).
CC	GN PROB OR PH4565.
CC	OS Pseudomonas aeruginosa.
CC	OC Pseudomonadaceae; Pseudomonas.
CC	NCBI_TAXID:287;
RN	SEQUENCE FROM N.A.
RC	STRAIN=ATCC 15692 / PA01;
RX	MEDLINE=882896162; PubMed=3399390;
RA	Hirashima A., Hirose T., Inayama S., Inokuchi Y., Jacobson A.B.; "Analysis of the complete nucleotide sequence of the group IV RNA coliphage SP.", Nucleic Acids Res. 16:605-621 (1988).
RT	-!- FUNCTION: Not yet known.
CC	-!- MISCLENNOUS: The readthrough protein A1 includes the coat protein sequence.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	DR PROB_PSEAE STANDARD; PRT; 372 AA.
CC	DR PIR; S01964; S01964.
CC	DR HSSP; P03615; IQBE.
CC	DR InterPro; IPR002703; Levi_coat.
CC	DR Pfam; PF0819; Levi_coat; 1.
CC	KW Coat protein.
CC	FT CHAIN 1 132 COAT PROTEIN.
CC	FT CHAIN 1 331 READTHROUGH PROTEIN A1.
SQ	SEQUENCE 331 AA; 36203 MW; 4A6642E4B52C6582 CRC64;
CC	Query Match Score 8.8%; Best Local Similarity 24.5%; Matches 39; Conservative 17; Mismatches 61; Indels 42; Gaps 6
CC	QY 37 LSFTGIDDMDTMAAAEQLFEATA DALYTFYD 152 DR 94 LSFTGSTDEEALARTE-LAALIPLVADLNDINPATVHNLASSGGCGNDNSPPDV 152
CC	DB 87 KTVEQQLKETQAFYLLGIIGRGEYDTEYAAQRARIGKIHLDVLGLG-PDVYLGA TYTGG 145
CC	QY 153 PVPDPVKPPD----GTGRYKCPAACY--RLGSIYEVGKEGSPDITY----- 191 DB 153 PVPDPVKPPD----GTGRYKCPAACY--RLGSIYEVGKEGSPDITY----- 191
CC	QY 146 DALADDVVADEGEEAARAVDFLPMKLTLFDQI 184 DR 192 ----ERGDEVSVTFDYALEDFLGNNTWNRNWQRL 221

RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.; "Complete genome sequence of <i>Pseudomonas aeruginosa</i> PAO1, an opportunistic pathogen.";	NCBI_TaxID=1148;
RT		RN [1] SEQUENCE FROM N.A.
RL	Nature 406:959-964 (2000).	RP MEDLINE=97061201; PubMed=8905231;
CC	-!- FUNCTION: Catalyzes the transfer of a phosphate group to glutamate to form glutamate 5'-phosphate which rapidly cyclizes to 5'-oxopropine.	RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirayama M., Sugura M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Murakami A., Nakazaki N., Naruo K., Okumura S., Shimpoo S., Takeuchi C., Wada T., Watanabe A., RA Yamada M., Yasuda M., Tabata S.;
CC	-!- CATALYTIC ACTIVITY: ATP + L-glutamate = ADP + L-glutamate 5'-phosphate.	RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.";
CC	-!- PATHWAY: Proline biosynthesis; first step.	RT DNA Res. 3:109-136 (1996).
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).	RT DNA Res. 3:109-136 (1996).
CC	-!- SIMILARITY: Belongs to the Glutamate 5-kinase family.	RL
CC	-!- SIMILARITY: Contains 1 PUA domain.	DNA Res. 3:109-136 (1996).
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).	CC
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).	CC
DR	EMBL; AE004870; AAC07953.1; -.	CC
DR	PIR; G83074; G83074.	CC
DR	HAMAP; MF_00456; ; 1.	CC
DR	InterPro; IPR001048; Aa_kinase.	CC
DR	InterPro; IPR001057; Glu_5kinase.	CC
DR	InterPro; IPR005715; Prob.	CC
DR	InterPro; IPR002478; PUA.	CC
DR	Pfam; PF00696; aakinase; 1.	CC
DR	DR PIR; PF01472; PUA; 1.	CC
DR	DR PRINTS; PR00474; Glu5Kinase.	CC
DR	SMART; SM00359; PUA; 1.	CC
DR	TIGRFAMS; TIGR00102; PROB; 1.	CC
DR	PROSITE; PS00902; GLUTAPROTE_5_KINASE; 1.	CC
DR	PROSITE; PS50890; PUA; 1.	CC
KW	KW Transferase; Kinase; Proline biosynthesis; Complete proteome.	CC
FT	DOMAIN 280 358 PUA.	CC
SQ	SEQUENCE 372 AA; 39845 MW; FAF2E81F6A8DEC36 CRC64;	CC
QM	3 NDNDTLYTAATYRGNGDGHALADRIGLDEAFIAWLSETFGIDDPTMAAALAEQPLFEATAD 62	CC
Db	138 NENDTIVVDTDIRG-----DNNTLAALYAN--LVEADL 169	CC
QY	63 ALVTD--FWDHLESYERTQDLFANSTKTVQLEKTAQEYLLGLGRGYDTEYAAQR-A 117	CC
Db	170 VILTDGRGMEDADPENNPDQLIYEARADDQLDAVAGSSAGALGRGGMOTKLRAARLA 229	CC
QY	118 RIGKIHDLVGLGPDPYLGAYTRYTYGLLDALADDVVA-----DRGEEARAAVDEL 167	CC
Db	230 RSG-----GHTIVG-----GRIEFLVDRURAGEPLGLLTLPDRSKAAR--KOW 272	CC
QY	168 VARFLPMLKLTFD 181	CC
Db	273 LAGHLQMGRGTLVLD 286	CC
RESULT 14		CC
ID	IF2_-SYNT3 STANDARD;	CC
AC	P72689;	CC
DT	01-NOV-1997 (Rel. 35, Created)	CC
DT	01-NOV-1997 (Rel. 35, Last sequence update)	CC
DT	10-OCT-2003 (Rel. 42, Last annotation update)	CC
DE	Translation initiation factor IF-2.	CC
GN	INFB OR SIRO744.	CC
OS	Synechocystis sp. (strain PCC 6803).	CC
OC	Bacteria; Cyanobacteria; Chroococcales; Synechocystis.	CC
OX	OX NCBI_TaxID=1148;	CC
RN	RN RP	CC
RT	RT SEQUENCE FROM N.A.	CC
RA	RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y., MIYAJIMA N., HIRAYAMA M., SUGURA M., SASAMOTO S., KIMURA T., HOSOUCHI T., MATSUNO A., MURAKAMI A., NARUO K., OKUMURA S., SHIMPO S., TAKEUCHI C., WADA T., WATANABE A., RA YAMADA M., YASUDA M., TABATA S.;	CC
RA	RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.";	CC
RA	RT DNA Res. 3:109-136 (1996).	CC
RA	RT "FUNCTION: One of the essential components for the initiation of protein synthesis. Protects its binding to the 30S ribosomal subunits. Also involved in the hydrolysis of GTP during the formation of the 70S ribosomal complex (By similarity)."	CC
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.	CC
CC	-!- SIMILARITY: Belongs to the IF-2 family.	CC
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sb-sib.ch).	CC
DR	DR EMBL; D90900; BAA16696.1; -.	CC
DR	DR PIR; S74544; S74544.;	CC
DR	DR HAMAP; MF_00100; ; 1.	CC
DR	DR InterPro; IPR0010795; EF_GTPbind.	CC
DR	DR InterPro; IPR004161; EFTU_D2.	CC
DR	DR InterPro; IPR0010178; IF2_.	CC
DR	DR InterPro; IPR001847; IF2_N.	CC
DR	DR InterPro; IPR005222; Small_GTP.	CC
DR	DR InterPro; IPR009000; Translat_factor.	CC
DR	DR Pfam; PF00009; GTP_BPTU; 1.	CC
DR	DR Pfam; PF03144; GTP_EFTU_D2; 2.	CC
DR	DR Pfam; PF04760; IF2_N; 2.	CC
DR	DR PRINTS; PR00315; EFTONGATNFCT.	CC
DR	DR PRODOM; PD18610; IF2; 1.	CC
DR	DR TIGRFAMS; TIGR00487; IF2; 1.	CC
DR	DR TIGRFAMS; TIGR00231; small_GTP; 1.	CC
DR	DR PROSITE; PS01176; IF2; 1.	CC
KW	KW Initiation factor; Protein biosynthesis; GTP-binding; Complete proteome.	CC
FT	FT DOMAIN 496 648 G-DOMAIN.	CC
FT	FT NP_BIND 502 509 GTP (BY SIMILARITY).	CC
FT	FT NP_BIND 552 556 GTP (BY SIMILARITY).	CC
FT	FT NP_BIND 606 609 GTP (BY SIMILARITY).	CC
SQ	SQ SEQUENCE 1001 AA; 108118 MW; 4C7F00DEE5CE31C CRC64;	CC
QY	QY Query Match Best Local Similarity 8.7%; Score 81; DB 1; Length 1001; Matches 47; Conservative 20; Missmatches 54; Indels 88; Gaps 7;	CC
Db	Db 22 LADRIGLDEFAETAWLSETFGIDDPTMAAALAEQPLFEATADALVTDYDHLEYERTQDL 81	CC
Db	Db 474 -AAAAKITTEMLDEADLNLVRRPPVVTIMGHVDHGKTTLDSTRKTYWAQGEAGGITOH 532	CC
QY	QY 106 GENDTE-----AVQITOTLDEETARMVAFEVAVETPERV-- 473	CC
Db	Db 423 LADLJKISTDIKRLFKRGV-----	CC
QY	QY 533 GATHVEVERNDKTEQIVELDTPHEAFTAMRANGARYVTDI 575	CC
QY	QY 146 DALADDVYADRGEEA----AAAVDELYA 169	CC
Db	Db 576 VVAADDGVQPQTKEAISHAKAAVGVLIV 604	CC

Best Local Similarity 21.4%; Pred. No. 47;
Matches 46; Conservative 31; Mismatches 77; Indels 61; Gaps 9;

Qy	5 NDTLVADVRNGIDSHA-----LADRIGDDEAIIAWRISFTGIDDDTMALAA 52
Db	552 DDAEAAADAMDALEMDAEMADGEIVDVIAIDQTMMLANASTEARARTGADGGA
Qy	53 E-QPFLFEATADAL-----VTDIYDHIESERTQDLFANSKTVQEQLKETQAE 98
Db	612 EVKTAEEESDAAEIDESRILLALQQVSVDADENRA--TSDTYSGRATVGDATALLDD 668
Qy	99 YLLGLGRGEYDTEYAAQR-----IGKIHDPVLGLGPDVYLGAATRYTGJL 145
Db	669 VVSFA---DTDTAGOTRAATDROHAASRVASAVDEVAGISETAAGA---TAVA 719
Qy	146 DALA-----DDVVADERGEEAAAVADELVAF 171
Db	720 DSAAQTQDTLSVDDAAIDLADR-ZAALDDLLAEEF 753

Search completed: August 10, 2004, 15:26:51
Job time : 6.79177 secs

RESULT 15

HTR2 HALSA	STANDARD;	PRT;	764 AA.
ID	HTR2 HALSA		
AC	P71410;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	Sensory rhodopsin II transducer (HTR-II) (Methyl-accepting phototaxis protein II) (MPP-II).		
DE	Halobacterium salinarium.		
OS	Archaea; Euryarchaeota; Halobacteria; Halobacte		
OC	Halobacteriaceae; Halobacterium.		
OX	Zhang W., Brooun A., Mueller M.M., Alam M.;		
RA	"The primary structures of the Archaeon Halobacterium salinarium blue light receptor sensory rhodopsin II and its transducer, a methyl-accepting protein."		
RT	Proc. Natl. Acad. Sci. U.S.A. 93:8230-8235(1996).		
RN	[2]		
RP	FUNCTION		
RX	PubMed=9515936;		
RA	Hou S., Brooun A., Yu H.S., Freitas T., Alam M.;		
RT	"Sensory rhodopsin II transducer HtrII is also responsible for serine chemotaxis in the archaeon Halobacterium salinarum."		
RL	Bacteriol. 180:1600-1602 (1998).		
CC	-!- FUNCTION: Transduces signals from the phototaxis receptor sensory rhodopsin II (SR-II) to the flagellar motor. Responds to light changes through the variation of the level of methylation. Also acts as a chemotransducer.		
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).		
CC	-!- SIMILARITY: Contains 1 methyl-accepting transducer domain.		
CC	-!- SIMILARITY: Contains 2 HAMP domains.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions. This statement is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).		
DR	EMBL; U62676; AAC44369.1; -		
DR	PIR; T44946; T44946.		
DR	HSSP; P02942; IQU7.		
DR	InterPro; IPR004089; Chmtaxis_transd.		
DR	InterPro; IPR004660; HAMP.		
DR	Pfam; PF00572; HAMP; 1.		
DR	Pfam; PF00015; MCPsignal; 1.		
DR	SMART; SM00304; HAMP; 2.		
DR	SMART; SM00283; MA; 1.		
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSUDC_2; 1.		
DR	PROSITE; PS50885; HAMP; 2.		
KW	Transducer; Photoreceptor; Chemotaxis; Transmembrane; Methylation; Repeat.		
FT	INIT_MBT 0 0 BY SIMILARITY.		
FT	DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).		
FT	TRANSMEM 16 36 POTENTIAL.		
FT	DOMAIN 37 277 EXTRACELLULAR (POTENTIAL).		
FT	TRANSMEM 278 298 POTENTIAL.		
FT	DOMAIN 299 764 CYTOPLASMIC (POTENTIAL).		
FT	DOMAIN 303 355 HAMP 1.		
FT	DOMAIN 397 450 HAMP 2.		
FT	DOMAIN 469 705 METHYL-ACCEPTING TRANSDUCER.		
SQ	SEQUENCEB 764 AA; 79187 MW; 1E0D7E460FC588 CRC64;		

Query Match 8.6t; Score 80.5; DB 1; Length 764;

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 10, 2004, 15:22:05 ; Search time 22.0533 Seconds
 (without alignments)
 2632.505 Million cell updates/sec

Title: US-09-455-978B-77

Perfect score: 933

Sequence: 1 MSNDNTLVTDVRNGIDGH.....DELVARFLPMLKULTFDQI 184

Scoring table: BLASTM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : SPTREMBL_25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp Rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archeap:*

ALIGNMENTS

17 92.5 9.9 503 16 Q8YDG6 brucella me
 18 92.5 9.9 514 2 Q8VOK3
 Q8dhh0 brucella ab
 19 90.5 9.7 194 16 Q8BH0
 20 87.5 9.6 400 16 Q82AV2
 Q82av2 streptomyce
 21 88.5 9.5 458 16 Q8PMR9
 Q8pmr9 xanthomonas
 22 88.5 9.5 505 11 Q99KK9
 Q99kk9 mus musciculi
 23 88.5 9.5 1361 16 Q8RHV1
 Q8rhv1 fusobacteri
 24 88 9.4 500 16 Q9TYD3
 Q9jyds neisseria m
 25 88 9.4 505 16 Q9JTB6
 Q9jtb6 neisseria m
 26 87.5 9.4 2246 2 Q9AKS6
 Q9aks6 pseudomonas
 27 86.5 9.3 157 17 Q9HNQ8
 Q9hnq8 halobacteri
 28 86.5 9.3 468 16 Q8p6N2
 Q8p6n2 xanthomonas
 29 86.5 9.3 895 11 Q8C117
 Q8c117 mus musciculi
 30 86 9.2 670 4 Q96CN5
 Q96cn5 homo sapien
 31 86 9.2 908 16 Q8CJKB
 Q8cjkb streptomyce
 32 85.5 9.2 965 16 Q81YS7
 Q81ys7 bacillus an
 33 85 9.1 306 16 Q8D7S6
 Q8d7s6 amycolatops
 34 85 9.1 2703 16 Q9KOTO
 Q9koto neisseria m
 35 84.5 9.1 266 17 Q9HND3
 Q9hnd3 halobacteri
 36 84.5 9.1 759 16 Q98C20
 Q98c20 rhizobium l
 37 84 9.0 170 3 Q871H7
 Q871h7 rhodopirell
 38 84 9.0 412 16 Q7UG88
 Q7ug88 rhodopirell
 39 84 9.0 421 2 Q9RQ25
 Q9rq25 amycolatops
 40 84 9.0 467 16 Q87DW5
 Q87dw5 xylolla fas
 41 84 9.0 491 16 Q92tQ4
 Q92tq4 rhizobium m
 42 83.5 8.9 258 16 Q82GU5
 Q82gu5 streptomyce
 43 83.5 8.9 712 10 Q80890
 Q80890 arabiopsis
 44 83.5 8.9 755 10 QBRX83
 Qbrx83 arabiopsis
 45 83.5 8.9 826 16 Q92TC2
 Q92tc2 rhizobium m

RESULTS

RESULT 1

Q9KEFH4 ID Q9KEFH4 PRELIMINARY; PRT; 439 AA.
 AC Q9KEFH4; DT 01-OCT-2000 (T-EMBLrel. 15, Created)
 RA Takanishi H.; Nakatone K.; Takaki Y.; Maeno G.; Sasaki R.; Masui N.;
 RA Fuji F.; Hirama C.; Nakamura Y.; Ogasawara N.; Kuhara S.;
 RA Horikoshi K.; RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis.";
 OC Bacteria; Firmicutes; Bactillales; Bactillaceae; Bacillace; Bacillus.
 OC NCBI_Taxid:86655;
 OC [1]
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAN= C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takanishi H.; Nakatone K.; Takaki Y.; Maeno G.; Sasaki R.; Masui N.;
 RA Fuji F.; Hirama C.; Nakamura Y.; Ogasawara N.; Kuhara S.;
 RA Horikoshi K.; RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:417-4331(2000).
 DR EMBL; AP001508; BAB04224; 1. .
 DR PIR; A83713; A83713.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0004871; F:signal transducer activity; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0006935; P:chemotaxis; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho . . IEA.
 DR GO; GO:0007165; P:signal transduction; IEA.
 DR InterPro; IPR004089; Chmataxis transl.
 DR InterPro; IPR002114; HPr_SerP_S.
 DR Pfam; PF00015; MCPSignal; 1.
 DR SMAR; SM00283; MA_1.
 DR PROSITE; PS50111; CHEMOTAXIS TRANSDUC_2; 1.
 DR PROSITE; PS00589; PTS_HPR_SER; 1.
 KW Complete proteome.
 SEQUENCE 439 AA; 48918 MW; 1E092BB2F42592AE CRC64;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	14.8	15..9	439 16 Q9KEFH4	O9kh4 bacillus ha
2	13.2	14..1	537 16 Q9AB06	Q9ab06 caulobacter
3	11.7..5	12..6	499 16 Q8UGL0	Q8ugl0 agrobacteri
4	10.5..5	11..3	555 16 Q9AY0	Q9ay0 caulobacter
5	10.2	10..9	434 16 Q814K6	Q814k6 bacillus ce
6	9.8..5	10..6	1222 16 Q8D4B1	Q8db1 vibrio vuln
7	9.7..5	10..5	248 10 Q8S4W2	Q8s4w2 phytopthor
8	9..7	10..4	218 16 Q88031	Q88031 streptomyce
9	9..5..5	10..2	1364 2 Q84GK0	Q84gk0 escherichia
10	9..5	10..2	433 16 Q81UN0	Q81un0 bacillus an
11	9..4..5	10..1	419 9 Q8W6U7	Q8w6u7 bacteriop
12	9..4..5	10..1	883 17 Q9HRW3	Q9hrw3 halobacteri
13	9..4	10..1	218 16 Q82MB8	Q82mb8 streptomyce
14	9..4	10..1	1090 10 Q644Z6	Q644z6 arabidopsis
15	9..3..5	10..0	514 16 Q8FUX2	Q8fux2 brucella su
16	9..3	10..0	881 16 Q8UJAL	Q8ujal agrobacteri

RESULT 4					
	ASY0	PRELIMINARY;	PRT;	555 AA.	
3	GQAQKDR-QLDB-----RLNPFGLGHGERQNQLSDMKGVTIGSDASIDRFTYKRAVPET	56			
79	QDLFPANSTKTVEQLKETQAEEYLGLGRGEYDTEYAAQRARIKGKHVDLGLGPDVYLGAYT	138			
57	AKFEFSSSEAHHTHAAEKSMLKHWRSRASGTFENEDTYNTAVTAISRHARISLEPFWYIGSA	115			
139	RYYTGGLDALAD	150			
116	IMLDGIVKAVIE	127			
SEQUENCE FROM N.A.					
STRAIN=ATCC 19089 / CB15;					
MEDLINE=21173698; PubMed=11259647;					
Nierman W.C., Feldblyum T.V., Lauten M.T., Paulsen I.T., Nelson K.E., Eisen J.J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddick J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phdke N.D., Ely B., DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haff D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O., Salzberg S.L., Ventler J.C., Shapilo L., Fraser C.M.;					
"Complete genome sequence of Caulobacter crescentus.";					
Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).					

RESULT 5						
Q814K6	ID	Q814K6;	PRELIMINARY;	PRT;	434 AA.	
AC	Q814K6;					
DT	01-JUN-2003	(TREMBLrel. 24, Created)				
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)				
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)				
DE	Methyl accepting chemotaxis protein.					
GN	BCS424.					
OS	Bacillus cereus (strain ATCC 14579 / DSM 31).					
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.					
OX	NCBI_TAXID=26900;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=2008415; PubMed=12721630;					
RA	Ivanova N., Sorokin A., Anderson I., Galleron N., Candelier B., Kapatral V., Bhattacharyya A., Resnik G., Mikhailova N., Lapidus A., Chu L., Mazur M., Goitsman E., Larsen N., D'Souza M., Walunas T., Grechuk Y., Pusch G., Haselkorn R., Fonstein M., Ehrlich S.D., Overbeek R., Kyriides N.,					
RA	"Genome sequence of <i>Bacillus cereus</i> and comparative analysis with <i>Bacillus anthracis</i> ."					
RL	Nature. 423:87-91(2003)					
DR	AE017015; AAP12286.1; -.					
DR	GO:0016020; C:membrane; IEA.					
DR	GO:0004871; P:signal transducer activity; IEA.					
DR	GO:0006935; P:chemotaxis; IEA.					
DR	GO:0007165; P:signal transduction; IEA.					
DR	InterPro; IPR004089; Chmtaxis_transd.					
DR	InterPro; IPR004090; Me chemotaxis.					
DR	Pfam; PF00015; MCPSignal_1.					
DR	PRINTS; PR00260; CHEMTRNSDUCR.					
DR	SMART; SM00283; MA_1.					
DR	PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.					
KW	Complete proteome.					
SQ	SEQUENCE 434 AA; 49441 MW; E1007BA02300205 CRC64;					
Qy	Query Match	10.9%	Score 102;	DB 16;	Length 434;	
Db	Best Local Similarity	20.5%	Pred. No. 1.7;			
Matches	34; Conservative	27;	Mismatches	71;	Indels 34; Gaps 3;	
Qy	31 AEIAURRLSFGIDDTMAAIAAEQPLFEATADAIVTDFDHLSEYERTQDLFANSTK--					87
Db	34 SELKIQMDMLHISKEQLQVVKVLQDFIYAEIDWTEKFY-----					-ANITKOPN 80
Qy	88 -----TVEBQKETOAEAYLGLGRGYDTEYAAQRARIGKHDVGLGPDVYLGYAT 138					
Db	81 LITITERYYSIPLKQTLTHIKELFSGNMNHEDIEQVRVRIAKRHVQIGLHLRKWYTAYQ 140					
Qy	139 RYYTGLDAADDVYADRSEEAAAVDELVARFLPMLKLTEDQI 184					
Db	141 ELFRSTIKLQTKI -----TTIDDEFSYINVINKLFTLBEQL 177					
RESULT 6						
Q8D4B1	ID	Q8D4B1	PRELIMINARY;	PRT;	1222 AA.	
AC	Q8D4B1;					
DT	01-MAR-2003	(TREMBLrel. 23, Created)				
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)				
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)				
DE	Conserved hypothetical protein.					
GN	VV1397.					
OS	Vibrio vulnificus.					
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibionales; Vibrionaceae; Vibrio.					
OX	NCBI_TAXID=672;					
RN	SEQUENCE FROM N.A.					
RP	STRAIN=CNCMP6					
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H., Chey H.E.; "Complete genome sequence of vibrio vulnificus Cemic-					

RL	Submitted (DBC-2002) to the EMBL/GenBank/DDBJ databases.	Db	217 SKLLK 221
DR	InterPro; IPR00437; Prok_1lipoprot_S.		
DR	Hypothetical protein; LIPOPROTEIN; 1.		
KW	Complete proteome.		
SEQUENCE	1222 AA; 136974 MW;		
Matches	37; Conservative 34; Mismatches 59; Indels 53; Gaps 7;		
Query Match	10.6%; Score 98.5; DB 16; Length 1222;		
Best Local Similarity	20.2%; Pred. No. 15;		
Matches	37; Conservative 34; Mismatches 59; Indels 53; Gaps 7;		
Qy	4 DDDTMALAAEAEQPLFEATDALVTDYDHLRSY-----ERTODLFANSTKTV 89		
Db	606 DSENTRILTEHOPIWSSPTDSSYEAYD--NNNNQKWENTRAYWERTTTELLADSSTKV 663		
Qy	90 -----EOLKETOAE-----YLIGLGRGEYDT-----EYAAQ 115		
Db	664 SGSSRYLIADSDAIIKRTSSEKPDGEKFSLEHTISATEKAYETWHIALEGYDETA 723		
Qy	116 RARICKHDLGLGPDVYLG-AYTRYTGGLDALADVVADRGEAAADELVARFLPM 174		
Db	724 ANVFGHQTYVENVLYDDIMGMGSYTQWDRNIDLAE--VDRLLQEVVPVDQITPENIPN 780		
Qy	175 LKL 177		
Db	781 LNL 783		
RESULT 7			
QBS4W2	PRELIMINARY;	PRT;	248 AA.
ID	QBS4W2;		
AC	QBS4W2;		
DT	01-JUN-2002 (TREMBLrel. 21; Created)		
DT	01-JUN-2002 (TREMBLrel. 21; Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25; Last annotation update)		
DE			
OS	Phytophthora infestans (Potato late blight fungus).		
OC	Bukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;		
OX	Phytophthora		
NCBI_TAXID	=4787;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=2197832; PubMed=11950882;		
RX	Dacks J.B., Doolittle W.F.;		
RX	"Novel syntexin gene sequences from Giardia, Trypanosoma and algae: implications for the ancient evolution of the eukaryotic endomembrane system".		
RX	J. Cell. Sci. 115:1635-1642 (2002).		
DR	InterPro: IPR000727; T_SNAKE.		
DR	Pfam: PF05739; SHARE_1.		
DR	SMART: SM00391; t_SNAKE; 1.		
DR	PROSITE: PS50192; T_SNAKE; 1.		
SQ	SEQUENCE 248 AA; 27934 MW;		
Query Match	10.5%; Score 97.5; DB 10; Length 248;		
Best Local Similarity	22.7%; Pred. No. 1, 9;		
Matches	42; Conservative 31; Mismatches 67; Indels 45; Gaps 7;		
Qy	6 DTLYTADVRNGIDGHALADRIGLDEAEI_AWRLSF-----TGIDDDTMAA-LA 51		
Db	68 ETIVMVEANRAKFPH-----IDAETASRKAFVAATRKELQAVSTIESTDIVKTR-RK 120		
Qy	52 AEQPLFEATDALVTDYDHLSEYRTQLEFASTKTVQQLKETQAEVLLGLRGELYDTE 111		
Db	121 EERKLMQPKSS-TSFRSNLTQGFRNERFLDEDTQRQQIMQFQNDISLAGL-----170		
Qy	112 YAAQRARIKHDVGLGPDVYLGATRYTGYGLDALLADDVADRGEAAADEVLARP 171		
Db	171 --- HSDITRLH----GVTEISSEVKHQNMKMDLTDV----DEAQERMFVMGR 216		
Qy	172 LPMLK 176		
RESULT 8			
OBS8031	PRELIMINARY;	PRT;	218 AA.
ID	OBS8031;		
AC	OBS8031;		
DT	01-NOV-1998 (TREMBLrel. 08; Created)		
DT	01-NOV-1998 (TREMBLrel. 08; Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25; Last annotation update)		
DE	Hypothetical protein SC06675.		
OS	Streptomyces coelicolor.		
OC	Bacterium; Actinobacteria; Streptomyctaceae; Streptomyctales;		
OC	Streptomyces coelicolor.		
NCBI_TAXID	=1902;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3 (2);		
RA	Seeger K.J., Harris D.; Rajandream M.A.;		
RA	Parkhill J., Barrell B.G., Rajandream M.A.;		
RA	Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.		
RL	Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3 (2);		
RA	Parkhill J., Barrell B.G., Rajandream M.A.;		
RA	Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3 (2);		
RX	MEDLINE=9700351; PubMed=8843436;		
RA	Redenbach M., Kieser H.M., Denapaire D., Eichner A., Cullum J.,		
RA	Kinashi H., Hopwood D.A.;		
RA	"A set of ordered cosmid and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3 (2) chromosome.".		
RT	MoL Microbiol. 21:77-96 (1996).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3 (2) / MI45;		
RX	MEDLINE=21996410; PubMed=12000953;		
RA	Bentley S.D., Chater K.F., Cerdeno-Taronga A.-M., Challis G.L.,		
RA	Thompson N.R., James K.D., Quail M.A., Kieser H.,		
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,		
RA	Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,		
RA	Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,		
RA	Rabinowitzsch E., Rajandream M.A., Rutherford K., Rutter S.,		
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,		
RA	Warren T., Wetzorek A., Woodward J., Barrell B.G., Parkhill J.,		
RA	Howwood D.A.;		
RA	"Complete genome sequence of the model actinomycete Streptomyces		
RT	coelicolor A3 (2).".		
RT	Nature 417:141-147 (2002).		
RL	EMBL; AL939128; CAA19954.1; -.		
DR	T35174; T35174.		
DR	GO:GO:0003677; F:DNA binding; IEA.		
DR	InterPro: IPR001387; DR:Protein binding; IEA.		
DR	PFAM: PF01381; HTH_3.		
DR	SMART: SM00510; HTH_XRE_1.		
DR	SRINE: SRINE_216.		
KW	Hypothetical protein; Complete Proteome.		
SQ	SEQUENCE 218 AA; 21670 MW;		
Query Match	10.4%; Score 97; DB 16; Length 218;		
Best Local Similarity	29.6%; Pred. No. 1, 7;		
Matches	53; Conservative 21; Mismatches 77; Indels 28; Gaps 10;		
Qy	7 TLVTDVNRNGIDGHALADRIGLDEAEI-AWRLSF-----LKETOAEYLLGLRGELYDTE 115		
Db	15 TLAAVARYG---ALADRLDPVAAEVFDVGRLSAASGPESVYAGLIGGRP---TGGAA 66		
Qy	64 LV-TDFYDLESYRTQLDIFANSTKTEQ-----LKEQAEYLLGLRGELYDTE 115		
Db	67 DVQTFRVQIDLRLRTR-LKPNGKTYQEIADAGMSRQQAGALIN-GDRRPPMEHCDA 124		
Qy	116 RARRGRKH-D-VLGDPVYLGATRYTGYGLDALLADDVADRGEAAADEVLARFL 172		

Db	125	IQRFFRVAHGFLTAEDPEALAGALQRTQQQL---ADRERQAAAPADDPLERLL	178	OX	NCBI_TaxID=198094;			
	[1]	SEQUENCE FROM N.A.		RN	SEQUENCE FROM N.A.			
	RN	MEDLINE=22508414; PubMed=12721629;		RX	RN			
	RP	Read T. D., Peterson S. N., Tourasse N., Baillie L.W., Paulsen I.T.,		RA	Read T. D., Peterson S. N., Tourasse N., Baillie L.W., Paulsen I.T.,			
	ID	Nelson K.E., Eisen H., Fouts D.E., Gill S.R.,		RA	Nelson K.E., Eisen H., Fouts D.E., Gill S.R.,			
	Q84GK0	PRELIMINARY;	PRT;	1364 AA.	Q84GK0	PRELIMINARY;	PRT;	1364 AA.
	ID	Holtzapfle E.K., Oksstad O.A., Helgason E., Ristone J., Wu M.,		RA	Holtzapfle E.K., Oksstad O.A., Helgason E., Ristone J., Wu M.,			
	AC	Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,		RA	Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.,			
	DT	DeBoy R.T., Madupu R., Daugherty S.C., Durkin A.S., Haft D.H.,		RA	DeBoy R.T., Madupu R., Daugherty S.C., Durkin A.S., Haft D.H.,			
	01-JUN-2003	(TREMBLrel. 24, Created)		RA	01-JUN-2003	(TREMBLrel. 24, Last sequence update)		
	DT	Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,		RA	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
	01-OCT-2003	Benton J.L., Mahamond Y., Jiang L., Hance J.R., Weidman J.P.,		RA	01-OCT-2003	Benton J.L., Mahamond Y., Jiang L., Hance J.R., Weidman J.P.,		
	DE	Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Niernan W.C.,		RA	DE	Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Niernan W.C.,		
	GN	Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salberg S.L.,		RA	GN	Hazen A., Cline R., Redmond C., Thwaite J.E., White O., Salberg S.L.,		
	OS	Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,		RA	OS	Thomason B., Friedlander A.M., Koehler T.M., Hanna P.C., Kolsto A.-B.,		
	OG	Fraser C.M.,		RA	OG	Fraser C.M.,		
	OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		RT	OC	"The genome sequence of <i>Bacillus anthracis Ames</i> and comparison to closely related bacteria.";		
	OX	Enterobacteriaceae; Escherichia.		RL	OX	Enterobacteriaceae; Escherichia.		
	RN	Plasmid pCS1.		DR	RN	Plasmid pCS1.		
	RP	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		DR	RP	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
	RC	NCBI_TaxID=562;		DR	RC	NCBI_TaxID=562;		
	SEQUENCE FROM N.A.			GO	SEQUENCE FROM N.A.			
	STRAIN=H10407;			GO	GO:0016020; C:membrane; IEA.			
	RA	Fleckenstein J.M., Patel S.K., Dotson J. ;		GO	GO:0004871; F:signal transducer activity; IEA.			
	RT	"Identification and molecular characterization of Bata, an autotransporter protein of enterotoxigenic Escherichia coli. ";		GO	GO:0006935; P:chemotaxis; IEA.			
	RT	Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.		GO	GO:0007165; P:signal transduction; IEA.			
	RL	EMBL:AY163491; AAC017297_1..		DR	InterPro: IPR040890; ChmotaXis transl.			
	DR	GO: GO:0046821; C:extrachromosomal DNA; IEA.		DR	InterPro: IPR04090; Me_chemoCtaxis			
	DR	GO: GO:0004252; F:serine-type endopeptidase activity; IEA.		DR	Pfam: PF00015; MCPsigna_1..			
	DR	GO: GO:0006508; P:proteolysis and peptidolysis; IEA.		DR	PRINTS; PR0026; CHEMTRNSDUCR.			
	DR	InterPro: IPR006315; Autotransport.		DR	SMART; SM00283; MA_1..			
	DR	InterPro: IPR005546; Autotransporter.		DR	PROSITE; PS50111; CHEMOTAXIS_TRANSduc_2..			
	DR	InterPro: IPR000710; Peptidase_S6.		KW	Complete proteome.			
	DR	InterPro: IPR004899; Peptidase_S6.		SQ	SEQUENCE 433 AA: 49391 MW: E06512BA0E696D92 CRC64;			
	DR	InterPro: IPR03797; Autotransporter_1.		Query	31 AEIAWRLSFTGIDDPTMAAALAAEOPFEATADALVTDYDHLESYERTQDFANSTK-- 87			
	DR	Pfam: PF02395; IgA1..1.		Db	33 SELKVQMDMLHISKEDLQIVKVLQFPIYBIDWTEKFY-----ANTKQPN 79			
	DR	Pfam: PF03212; Peractin_1..		Query	88 -----TVEQIKETOAQEYLIGGRGEYDTEAAQRARIGKIHDLVGLGPDVLYGAYT 138			
	DR	InterPro: PR00921; IGASERPFASE.		Db	80 LITIITERSSSIPKLUOTLKTHIKEFSGDMHDEFLEQVTKIAKTHVQIGHRKWYTAAYQ 139			
	DR	TIGRFAMs; TIGR01414; autotrans_bar1..1.		Qy	139 RYYTGLDLADDVVAADRGEAAAADVDELVARFELPMKLTFDQO 184			
	KW	Plasmid.		Db	140 ELFRSIMKILKTKI-----TTIDDESYSINVINKLFTLEQEL 176			
	SEQUENCE FROM N.A.			DE	4..			
	STRAIN=H10407;			OS	Bacteriophage phiE125.			
	RA	Fleckenstein J.M., Patel S.K., Dotson J. ;		OC	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;			
	RT	"Identification and molecular characterization of Bata, an autotransporter protein of enterotoxigenic Escherichia coli. ";		OC	Lambda-like viruses.			
	RT	Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.		OX	NCBI_TaxID=180504;			
	RL	InterPro: IPR047491; ANL40277..		RN	SEQUENCE FROM N.A.			
	DR	InterPro: IPR006944; Phage portal.		RA	Woods D.E., Jeddoh J.A., Fritz D.F., Deshazer D.;			
	DR	InterPro: IPR006427; Porta_HK97.		RT	"Burkholderia thailandensis B125 Harbors a Temperate Bacteriophage Specific for Burkholderia mallei."			
	DR	InterPro: IPR006427; Porta_HK97.		RL	Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.			
	DR	InterPro: IPR006427; Porta_HK97.		DR	AF447491; ANL40277..			
	DR	InterPro: IPR006427; Porta_HK97.		DR	InterPro: IPR006944; Phage portal.			
	DR	InterPro: IPR006427; Porta_HK97.		DR	InterPro: IPR006427; Porta_HK97.			
	Pfam: PF04860; Phage_portal..			DR	Pfam: PF04860; Phage_portal..			

QY	10.1%; Score 94.5; DB 9; Length 419;	Db	180 DVKSNVEGOLDRLADQIA-DKEDADPHDRSLINTAALEVTAJIEHPPAEREQARQTRD 237
Query Match Sequence	419 AA; 46404 MW; C15CC08F21D2C1 CRC64;	QY	63 ALVTDYDLESVERTQDLFANSTKTVBOLKETQAELYLLGRGEYDETEYAAQRA-----117
Best Local Similarity	21.7%; Pred. No. 7.2;	Db	238 ---DAADVLERFEERTDALADETTLADYREAVE-----AERETLDAEVSIDHRE 286
Matches	49; Conservative 28; Mismatches 74; Indels 75; Gaps 10;	QY	118 RIGHKHVD-----LGL-GPDVYLGAYTRYTGLLDLAD--DVADRGEEAAAVDEL 167
QY	5 NDTLVADYRNGIDGHALADRIGLDEAEIAWRLSFTGIDDDTMALAAB----QPLFEAT 60	Db	287 RASDIDDEAAALAADDGLDDPDAEASAER-----DAVADOREAAAVREVAPAVSRL 340
DDB	201. NGTALGVIERPKDAPALDKQASVDRITDWNKAEGSSNAKKVALLOFGMTFRPLSMTN 260		
61 ADALNTD-----FYDHT-----SYE 76			
QY	77 RTQDLFANSRKTVBOLKETQAELYLLGRGEYDETEYAAQRAARTGK----IHFDVLGE- 128	RESULT 13	Q82NBB8 PRELIMINARY; PRT; 218 AA.
Db	261 VDAALIDALRLSALDIARYKIPAHMNVNEELERATSNTIHQSLQFVITYLPPWKRHEA 320	Q82NBB8 ID; PRELIMINARY; PRT; 218 AA.	
QY	321 KTRDLILPER----KQFIEFYIAGLRLGQDSSRYAA-YAVGRQMGWLISINDRRE 373	AC	Q82NBB8; PRELIMINARY; PRT; 218 AA.
Db	77 RTQDLFANSRKTVBOLKETQAELYLLGRGEYDETEYAAQRAARTGK----IHFDVLGE- 128	DT	01-JUN-2003 (TREMBlrel. 24, Created)
QY	129 -----GPDVYLGAYTRYTGLLDA-LADDVADRGEEAAAVDEL 167	DT	01-OCT-2003 (TREMBlrel. 24, Last sequence update)
Db	374 NMPPVKGGDIYLSP----NMVMDASKPQQLPVKGSEPPKAIDEI 414	DT	01-OCT-2003 (TREMBlrel. 25, Last annotation update)
		DE	Putative transcriptional regulator.
		GN	SAV742.
		OS	Streptomyces avermitilis.
		OC	Bacteria; Actinobacteridae; Streptomyctaceae; Actinomycetales; Streptomyces; Streptomyces; Streptomyctaceae; Streptomyces.
		OX	NCBI_TaxID:33903; NCBI_TaxID:1.
		RN	SEQUENCE FROM N.A.
		RP	SEQUENCE FROM N.A.
		RC	STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
		RX	MEDLINE=21477403; PubMed=11527948;
		RA	Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C., Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Onoae T., Kikuchi H., Shiba T., Sakai Y., Hattori M., Omura S.; "Genome sequence of an industrial microorganism Streptomyces avermitilis: deducing the ability of producing secondary metabolites.", Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
		RA	Shinose M., Hattori M., Omura S.; "Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis.", RIBI Biotechnol. 21:526-531 (2003).
		RT	EMBL: AP005028; BAC69453.1; GO: GO:0003677; F:DNA binding; IEA.
		DR	InterPro: IPR01387; HTH_3.1; Pfam: PF01381; HTH_3.1; SMART: SM00530; HTH_XRE; 1.
		KW	Complete proteome.
		SQ	SEQUENCE 218 AA; 24060 MW; 3227B89A840B4C73 CRC64;
			Query Match 10.1%; Score 94; DB 16; Length 218;
			Best Local Similarity 28.7%; Pred. No. 3.1; Mismatches 73; Indels 22; Gaps 8; Matches 47; Conservative 22; Mismatches 47; Score 94; DB 16; Length 218;
		QY	21 ALADRIGLDEAEI-AWRLSF-TGIDDDTMALAABQPLFFRATA DALVTDYDHLESYER 77
		Db	25 ALADRIGVVAHEVFVGRLSYASGPFPVKKLSSGRPAGEPDQAA-RFLQRDLRR 81
		QY	78 TDQLFANSRKTVQE-----LKETDAEYLIGLGRGEYDETEYAAQRAARIGKHD--VLGL 128
		Db	82 TR-LKPNNRRRTQQEADGAGMSRQQAGALIN-GDRPTMEHCDATQFRVHAGFLTAE 139
		QY	129 GPDVYLGAYTRYTGLLDA-LADDVADRGEEAAAVDEL 172
		Db	140 DPEALACTLQSRSEQELIQQL---ADRERAAMAVDDPLERLL 178
		RESULT 14	O64476 PRELIMINARY;
		ID	1090 AA.

AC 064476; 1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update).
Paratrite DNA2-NAM7 helicase family protein.
DE AT2019120.
GN OS Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids;
euroids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TAXID=3702;
RN _[1].
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20082487; PubMed=10617197;
RA Lin X., Raul S., Rounseley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii Y., Mason T.M., Bowman C.L., Barnstead M.E., Felbliyum T.V.,
RA Bueli C.R., Kethum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Coppenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.,
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana"; Nature 402:761-768(1999).
RN _[2].
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
Lin X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AC002392; T00533.
DR GO; GO:0004386; F1 helicase activity; IEA.
KW Helicase.
SQ SEQUENCE 1090 AA; 121522 MW; EB170342E18DCA1F CRC64;
Query Match Score 94; DB 10; Length 1090;
Best Local Similarity 26.9%; Pred. No. 31;
Matches 39; Conservative 17; Mismatches 55; Indels 34; Gaps 5;
Job time : 24.0533 secs

Qy 55 PLFE---ATADALVTDYDHLESVERTQDLFA-----NISTKIVQEQLKETOAEYLGL 103
Db 441 PSYEETPAPDCFTTSFVDLHRSRNPQLIAHWAAMMHTAAGTSSGVKK-QEPWPFTL 499
Qy 104 GRGEYDTEYAAQRARIGKHDWVGLGRDVLXGAYTRYTGHDATA-----D 150
Db 500 VQGPPTG-----GKTHTVGMUNVHLVQYQQYTSLIKKLAPETYQNANECSSD 550
BRA1050.

Qy 151 DVVADRGEEAAAVADELVAREFLPML 175
Db 551 NILSGSDEVQMDHNLFPLKL 575

RESULT 15
Q8FDR2 PRELIMINARY; PRT; 514 AA.
ID Q8FDR2
AC 08FDR2;
DT 01-MAR-2003 (T-TREMBLrel. 23, Created)
DT 01-MAR-2003 (T-TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-TREMBLrel. 25, Last annotation update)
DE Peptide ABC transporter, periplasmic peptide-binding protein.
GN OS *Bacillus suis*
OC Bacteria; Proteobacteria; Alpha proteobacteria; Rhizobiales;
OC *Bacillus suis*; *Bacillus*
NCBI_TAXID=29461;
RN _[1].
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
Daugherty S.C., Debey R.T., Durkin A.S., Kolenow J.F., Madupu R.,

3
101
Art